REMARKS

The Office Action

The specification is objected to for informalities.

Claims 18-22 and 27-32 are pending in this application. All pending claims stand rejected under 35 U.S.C. § 112, first paragraph, for both inadequate written description and lack of enablement, and under 35 U.S.C. § 112, second paragraph, for indefiniteness.

Objection to the Specification

The legend of Figure 2 has been amended in accordance with the Examiner's suggestion. This objection may now be withdrawn.

Rejections Under 35 U.S.C. § 112, first paragraph

All pending claims stand rejected under 35 U.S.C. § 112, first paragraph, for both inadequate written description and lack of enablement. Each aspect of this rejection will be addressed separately.

Written Description

In rejecting the pending claims, the Examiner asserts that the invention encompasses the "administration [of] peptides that are allelic variants or corresponding sequences from other species, mutated sequences, splice variants or wild-type sequences as encompassed by the generic recitation of FREAC3" and that "[t]he instant disclosure of a single polypeptide, that of SEQ ID NO:2 does not adequately support the scope of the claimed FREAC3 genus" (Office Action mailed March 24, 2003, page 3, lines 13-20). As noted by the Examiner, the Federal Circuit has stated that a genus may be described either by a representative number of species or by a recitation of structural features common to the genus, which features constitute a substantial portion of the genus. *Regents of the University of California v. Eli Lilly & Co.* 119 F.3d 1559, 43 U.S.P.Q.2d

1398 (Fed. Cir. 1997). Applicants respectfully submit that the present specification adequately describes the invention as claimed by describing both a representative number of species and by a recitation of features common to the genus.

Contrary to the Examiner's assertion, Applicants describe not one, but six variants of the FREAC3 gene and protein isolated from human subjects. The protein sequence of SEQ ID NO:2 is the prototypical or "wild-type" FREAC3 protein most commonly observed in Applicants' studies. In addition, Applicants describe two allelic variants which retain FREAC3 biological activity. These variants have a GGC codon insertion following nucleotides 1122 or 1338 of SEQ ID NO:1, giving rise to a glycine insertion at amino acid positions 375 and 447, respectively. Applicants also describe three allelic variants that are biologically inactive including Ser82Thr substitution caused by G245C mutation, Ile87Met substitution cause by C261G mutation, and the deletion of the genomic sequence corresponding to nucleotides 93-102 of the cDNA (see, for example, page 3, line 23 through page 4, line 7, and page 45, lines 5-9, of the specification). Thus, by identifying six allelic variants of the human FREAC3 gene and its product, Applicants have disclosed a representative number of members falling within the genus of human FREAC3 proteins encompassed by the claims as presently amended. Applicants respectfully submit that, for this reason alone, the written description requirement as set forth by the *Lilly* court is satisfied and the rejection should be withdrawn.

In addition to providing a representative number of human FREAC3 variants, Applicants also describe FREAC3 using structural features common to the genus. For example, Applicants disclose that FREAC3 is a member of the forkhead family of transcription factors which a skilled artisan at the time of application filing would immediately understand to contain a characteristic and conserved DNA binding domain (see, for example, Kaufmann *et al. Mechanisms of Development*, 57: 3-20, 1996; copy enclosed). Applicants identify the forkhead domain at amino acids 69-178 (specification at page 43, line 10, and Figure 2 – boxed sequence). The criticality of the forkhead DNA

binding domain is demonstrated by Applicants use of molecular epidemiology. Applicants identified single amino acid substitutions (Ser82Thr and Ile87Met) in the DNA binding domain of human patients having clinical disease (specification at page 19, lines 20-21). Applicants further teach that the serine residue corresponding to Ser82 of FREAC3 is invariant in >80 forkhead family members in species ranging from yeast to human (specification at page 44, lines 1-2). Also, the isoleucine corresponding to Ile87 of FREAC3 is conserved in 88% of forkhead proteins and, in the remaining 12%, has never been reported as a methionine (specification at page 44, lines 14-16). Thus, the combination of Applicants' disclosure and information available in the art at the time of filing regarding the structure of forkhead transcription factors is sufficient to describe the FREAC3 molecules that have 95% identity to SEQ ID NO:2, are recited in the presently amended claims.

Applicants respectfully submit that the specification provides adequate written description of the FREAC3 polypeptides that fall within the claims. Applicants note, as detailed above, that the genus of FREAC3 polypeptides is described both by a recitation of an adequate number of species and by a recitation of structural features common to the genus; either one being sufficient to satisfy the written description requirement. Accordingly, Applicants respectfully request that this rejection be withdrawn.

Enablement

All pending claims stand further rejected under 35 U.S.C. § 112, first paragraph, for lack of enablement. Specifically, the Examiner asserts that the specification

> does not reasonably provide enablement for a method of increasing the FREAC3 biological activity in a mammal diagnosed as having developmental defect or a disease of the eye via administration of any particular compound or via any particular polypeptide (Office Action mailed March 24, 2003, paragraph spanning pages 5 and 6).

Although Applicants do not agree with the Examiner's assertion, in order to expedite prosecution, the claims have been amended to recite the treatment of a developmental defect or a disease of the via by administering a biologically active FREAC3 polypeptide having 95% sequence identity to SEQ ID NO:2. Thus, the claims no longer broadly recite increasing FREAC3 biological activity by any means and the specification clearly enables the full scope of the claimed method.

Applicants note that human testing is not required, for enablement purposes, to support claims of an *in vivo* utility. The Federal Circuit has repeatedly stated the following:

> Title 35 does not demand that such human testing occur within the confines of Patent and Trademark Office (PTO) proceedings ... Congress has given the responsibility to the FDA, not to the [PTO], to determine ... whether drugs are sufficiently safe. *Scott v. Finney*, 34 F.3d 1058, 1063 (Fed. Cir. 1994), (affirming *In re Watson*, 517 F.2d 465, 476 (C.C.P.A. 1975) and *In re Sichert*, 566 F.2d 1154, 1160 (C.C.P.A. 1977)).

To further elaborate this point, the first paragraph of § 112 "requires nothing more than objective enablement" and, in a case in which the Patent Office questions the enablement of a claim: "... it is incumbent upon the Patent Office ... to explain why it doubts the truth or accuracy of any statement in a supporting disclosure and to back up assertions of its own with acceptable evidence or reasoning which is inconsistent with the contested statement." *In re Marzocchi*, 439 F.2d 220, 169 U.S.P.Q. 367 (C.C.P.A. 1971).

Applicants respectfully submit that the specification enables the full scope of the invention as presently claimed, and that assertions to the contrary are, to date, insufficiently supported. Accordingly, this rejection may now be withdrawn.

Rejections Under 35 U.S.C. § 112, second paragraph

All pending claims stand rejected under 35 U.S.C. § 112, second paragraph, for indefiniteness. Specifically, the Examiner asserts that the claim terms "FREAC3," "wild-type FREAC3," and "FREAC3 biological activity" are indefinite. Each of these bases of rejection is separately addressed.

Applicants respectfully submit that claim term "FREAC3," as used in the claims, is not indefinite. As used in the present specification, FREAC3 refers, collectively, to all allelic variants of the human FREAC3 gene and gene products. Applicants provide the wild-type FREAC3 cDNA and polypeptide sequence (SEQ ID NOs: 1 and 2, respectively) and also disclose two allelic variants that do not affect biological activity (identified in phenotypically normal individuals) and three biologically inactive variants (identified in patients having clinical disease). The biologically active variants contain a GGC insertion following nucleotide 1122 or nucleotide 1338 of the FREAC3 cDNA, resulting in functional FREAC3 proteins having a glycine insertion at amino acid position 375 or 447, respectively (see, for example, specification, page 45, lines 5-9). The inactive variants are (i) Ser82Thr substitution caused by G245C mutation, (ii) Ile87Met substitution cause by C261G mutation, and (iii) the deletion of the genomic sequence corresponding to nucleotides 93-102 of the cDNA (see, for example, page 3, line 23 through page 4, line 7, and page 45, lines 5-9). Furthermore, human FREAC3 was previously identified in the art and recognized as a member of the forkhead protein family (see, for example, Kaufmann *et al. Mechanisms of Development*, 57: 3-20, 1996; copy enclosed).

Applicants provide the cDNA and amino acid sequence a total of six allelic variants of the human FREAC3 gene; three of which are biologically active and three of which are not. Thus, when the specification is read in it's entirety, and in view of the relevant prior art, the claim term "FREAC3" is clearly defined.

Likewise, Applicants submit that the term "wild-type FREAC3" is also clearly

defined as referring to the "amino acid sequence most often observed among members of a given animal species and not associated with a disease phenotype" (Specification at page 8, lines 19-22). However, in order to expedite prosecution, the claims have been amended to remove this term.

Finally, Applicants also submit that the term "FREAC3 biological activity" is not indefinite. The specification discloses that FREAC3 is a member of the forkhead family of transcription factors (specification at page 2, lines 17-19), that it binds to the DNA consensus sites of SEQ ID NOs: 3 and 4, and causes transcriptional activation. In this regard, Applicants provide several assays for detecting and defining FREAC3 biological activity. The clear tests for FREAC3 biological activity include reporter gene assays in which the FREAC3 consensus site is operably linked to lacZ or green fluorescent protein (specification, page 30, lines 19-25), ELISA assays (page 32, line 20 through page 33, line 9), and DNA binding assays (page 33, line 19 through page 34, line 9). Furthermore, and as noted previously, the prior art provides other relevant additional information for defining "FREAC3 biological activity" (see, for example, Kaufmann *et al. Mechanisms of Development*, 57: 3-20, 1996; copy enclosed). Thus, Applicants submit that the specification clearly defines "FREAC3 biological activity" in a definite manner.

In sum, the specification allows the skilled artisan to recognize, test, make, and use FREAC3 from the disclosure, both in terms of FREAC3 structure and FREAC3 biological activity. The skilled artisan can easily discern the metes and bounds of the claims.

This rejection should be withdrawn and such action is respectfully requested.

CONCLUSION

Applicants submit that the claims are in condition for allowance, and such action is requested. Enclosed is a petition to extend the period for replying for three months, to and including September 25, 2003. If there are any charges or any credits, please apply them to Deposit Account No. 03-2095.

Respectfully submitted,

Date: September 16, 2003

Kristina Bieker-Brady, Ph.D.
Reg. No. 39,109

Clark & Elbing LLP
101 Federal Street
Boston, MA 02110
Telephone: 617-428-0200
Facsimile: 617-428-7045

ELSEVIER

Mechanisms of Development 57 (1996) 3–20

Review article

Five years on the wings of fork head

Eckhard Kaufmann*, Walter Knöchel

Abteilung Biochemie, Universität Ulm, Albert-Einstein-Allee 11, 89081 Ulm, Germany

Received 12 February 1996; revision received 9 April 1996; accepted 9 April 1996

Abstract

Since its discovery five years ago the conserved family of fork head/HNF-3-related transcription factors has gained increasing importance for the analysis of gene regulatory mechanisms during embryonic development and in differentiated cells. Different members of this family, which is defined by a conserved 110 amino acid residues encompassing DNA binding domain of winged helix structure, serve as regulatory keys in embryogenesis, in tumorigenesis or in the maintenance of differentiated cell states. The purpose of this review is to summarize the accumulating amount of data on structure, expression and function of fork head/HNF-3-related transcription factors.

Keywords: Fork head; HNF-3; Transcription factor

1. Introduction

Transcription factors are usually classified according to the structural motif involved in DNA binding or, in some instances, to the oligomerization domain (Ptashne, 1988; Mitchell and Tjian, 1989; Ptashne and Ghann, 1990; Tjian and Maniatis; 1994; Nelson et al., 1995). Well known examples are the helix-turn helix motif found in homeobox-containing proteins, the paired box, the POU domain, the C2/H2 as well as the C2/C2 zinc finger motif, the helix-loop-helix and the leucine zipper motifs. An additional motif, the fork head/HNF-3 motif, was detected five years ago in the *Drosophila* fork head gene product and in rat hepatocyte nuclear factors 3 (HNF-3).

The *Drosophila* fork head gene is essential for the proper formation of terminal structures of the embryo. Mutations of the gene cause a homeotic transformation of gut structures into head-derived elements. In situ localisation studies revealed the presence of mRNA and protein at early stages predominantly in tissues derived from the endoderm germ layer. The nuclear localization of the protein led to the suggestion that it may function as transcriptional regulator (Weigel et al., 1989). Shortly after the description of fork head, a small family of hepatocyte-enriched DNA-binding transcription factors in rodents, the HNF-3 protein family, was described (Lai et al., 1990, 1991). The DNA-binding domain of the HNF-3 factors encompasses a region of about 110 amino acid residues and, at the time of its discovery, the sequence did not resemble any known DNA binding motif (Lai et al., 1990). However, by visual amino acid sequence comparison a high degree of sequence identity was found within the DNA binding domains of HNF-3α and the *Drosophila* protein fork head (Weigel and Jäckle, 1990). This domain, the fork head/HNF-3 domain, was later detected within many proteins of different species ranging from yeast to human (Lai et al., 1993). Thus, fork head and HNF-3 have to be considered as the founder members of a family of transcription factors which for structural reasons are also called winged helix proteins.

2. Structure and DNA contact

The structure of the fork head/HNF-3 domain has been solved by X-ray crystallographic analysis on the complex of the fork head domain of HNF-3γ with its cognate DNA at 0.25 nm resolution (Clark et al., 1993). The DNA target

* Corresponding author. Tel.: +49 731 5023288; fax: +49 731 5023277.

PII: S0925-4773(96)00539-4

site for HNF-3γ employed in the diffraction study includes 11 residues from the 5′ promoter region of the transthyretin (TTR) gene. The amino-terminal region of the domain is dominated by a cluster of three α-helices (H1, H2, H3). A loop region (T′) encompassing eight residues is located between H2 and H3. A β-strand (S1) between H1 and H2 and two further β-strands (S2, S3) which are separated by a non-ordered wing-like region W1 are located just behind H3 towards the carboxyl-terminus. All three β-strands form a three-stranded, antiparallel β-sheet. Finally, a second non-ordered wing-like structure W2 is found behind S3, just at the carboxyl-terminus of the domain. The association of H3 with S2 is further strengthened by the coordinate binding of a magnesium ion which seems to be essential for proper folding. Mutations of amino acids being engaged in this coordination completely abolish DNA binding (Clevidence et al., 1993; Häcker et al., 1995). In conclusion, the mutual interaction of the various secondary structure elements leads to a three-dimensional structural entity, resembling the shape of a butterfly: a core derived from the α-helical and β-sheet elements is flanked by two 'wings', W1 and W2. Therefore, this domain is synonymously called the 'winged helix' structure (Clark et al., 1993). It represents a structural unit which cannot be split further without losing DNA-binding properties, as revealed by footprinting and deletion studies (Lai et al., 1990, 1991; Kaufmann et al., 1994).

The principal base contacts with the major groove of the 13° bent TTR promoter DNA are mediated by the recognition helix H3. The only side chain minor groove base contact is within W2. Five interactions are due to elements other than W2 and H3. Numerous interactions with the sugar-phosphate backbone are made by the other structural elements leading to a total number of 14 contacts. This high number approaches the contacts made by dimeric helix-turn-helix (HTH) proteins or by other homeoproteins and thus ensures a high binding specificity (Brennan, 1993). A previous comparative analysis of a region related to the binding domain of HNF-3γ differs from the X-ray results since it does not include all β-strand elements and does not account for contacts with the DNA (Li and Tucker, 1993). A comparison of several winged helix domains shows that 25% of the amino acid residues are absolutely conserved. Mutations occur on surface residues or by conservative exchanges of buried residues. Therefore, all fork head domains involved in DNA binding should have a similar three-dimensional structure (Clark et al., 1993).

A relationship between the structures of the H2-T′-H3 motif in the fork head/HNF-3 domain and the HTH domains has been discussed (Brennan, 1993). The HTH motif is a widespread DNA-protein recognition motif (Schultz et al., 1991; Wolberger et al., 1993). Prokaryotic HTH proteins ensure specific major groove DNA contacts by binding as dimers and the inclusion of more distal residues, whereas eukaryotic HTH proteins, such as the homeodomain proteins, bind mostly as monomers and include distant basic residues for extra minor groove contacts (RXR motif). However, the binding of eukaryotic HTH proteins without such a second 'anchor' motif is not well understood (Brennan, 1993). Therefore, the solution of the structure of the fork head/HNF-3 domain has contributed to our understanding of the binding of this 'anchor-less' class of proteins.

A number of proteins which are only distantly related to the fork head domain in terms of their primary sequence show an astonishing structural similarity to the winged helix domain. Histone H5 which is involved in the formation of the 30 nm filaments can be easily superimposed on the structure of the HNF-3γ fork head domain (Clark et al., 1993). The root mean square of structural equivalent α-C atoms is only 0.13 nm, whereas the primary amino acid sequence similarity is only 12%. The only apparent difference between the two structures is within the W2 region which is not present in histone H5 (Ramakrishnan et al., 1993). This structural similarity does not seem to be purely accidental, but might imply that HNF-3 proteins could be involved in positioning of nucleosomes as shown by studies on the remote enhancer element of the albumin promoter (McPherson et al., 1993). Furthermore, additional unexpected similarities of the 3-D fold of other proteins with the fork head domain are reported. The DNA binding domain of the interferon regulatory factor 2 (Uegaki et al., 1995), the LexA repressor DNA binding domain (Fogh et al., 1994), the *Drosophila* heat shock factor dHSF (Vuister et al., 1994) and the ets-domain from human Fli-1 and murine Ets-1 strongly resemble the winged helix structure (Liang et al., 1994; Donaldson et al., 1996). If the definition of HTH is loosened in such a manner that variations with respect to the loop or turn region are permitted, all these proteins are then classified as HTH-related entities. Referring to the observed similarities one might speculate on the functional implications involved (Brennan, 1993; Clark et al., 1993).

3. Sequences of fork head-related genes from different organisms

Since the discovery of the winged helix domain in the *Drosophila* fork head protein and the rodent HNF-3 factors more than 80 genes encoding this evolutionary conserved domain have been identified. Most of these genes have been isolated by their DNA sequence similarity with *Drosophila* fork head/rat HNF-3, such as the FD family of fork head-related proteins in *Drosophila* (Häcker et al., 1992), the HFH family in rodents (Clevidence et al., 1993, 1994; Hackett et al., 1995), MF-1/MF-2 and the fkh family in mouse (Kaestner et al., 1993; Sasaki and Hogan, 1993) as well as the XFD family of *Xenopus* fork head domain proteins (Knöchel et al., 1992, Scheucher et al.,

1995). Other genes encoding fork head proteins have been isolated by genetic screens (HCM1 in yeast: Zhu et al., 1993), by promoter trapping (PES-1 in *Caenorhabditis elegans*: Hope, 1994), by P-element-mediated enhancer trapping (slp-1 and slp-2 in fly: Grossniklaus et al., 1992), by affinity chromatography (SGF1 in *Bombyx mori*: Mach et al., 1995) or by screening of expression libraries with known target sites (human ILF-1/2: Li et al., 1992a; MNF in mouse: Bassel-Duby et al., 1994).

Table 1 shows a computer-supported alignment of the published winged helix domain sequences (we are aware of a large number of additional genes being under current investigation, but here we only refer to work published up to the end of 1995). Although some structural elements are common to all listed sequences, the overall sequence identities exhibit great variability (between 100 and 30%) with respect to the fork head domain of fkh. Conserved regions are found in front of and in the H1, H2 and H3 regions which mediate most direct base/protein contacts and, as basic residues, at the carboxyl-terminus. We found that, due to conserved residues at distinct positions in the individual sequences, the different fork head domains can be classified into different subgroups. We suggest a scheme of ten different classes, some of which can be further subdivided into sub-classes. In this scheme, class 1 (fkh/HNF-3) is characterized by A(9), L(43), Q(51), N(92) and C(98), a first sub-class is defined by T(7) and F(46), a second by F(2) and P(7), a third by N(5), S(7), M(31) and E(76, 89), and a fourth by A(7), A(85), and E(110). Although such sub-classes are chosen somewhat arbitrarily (e.g., fkh and SGF-1 may define another one), they may prove quite useful for detecting species homologues or for evaluating common features of individual genes (genomic structure, binding motifs, etc.). Based on the classification outlined in Table 1, individual fork head domains can be unambiguously identified by suitable choice of a maximum of five identifier residues (colour code in Table 1).

In contrast to their high similarity within the DNA binding domain, only a few winged helix proteins are known to be homologous with respect to their entire sequence. This is partially due to the fact that in many cases only the DNA binding domain but not the flanking regions have been reported (complete sequences are marked in Table 1 by asterisks). The three mouse HNF-3 proteins are the true interspecies homologues to the respective proteins in rat. *Xenopus* fork head proteins XFD-3/3′ are the amphibian counterparts to rodent HNF-3β, whereas XFD-7/7′ correspond to HNF-3α. Mouse HFH-4 corresponds to rat HFH-4. Rat BF-1 corresponds to human HFK1 and avian qin, and mouse MFH-1 is closely related to *Xenopus* XFD-4. Although flanking sequence information is not available, in many other cases a very close relationship is, at least, very likely. Members of one individual class (Table 1) do often exhibit identity rates of 95–100%. Based on the sequence similarity it is reasonable to assume that, for example, human FREAC-3 and mouse fkh-1/MF-1, FREAC-4 and rat BF-2/HFH-B2, FREAC-6, rat HFH-3 and *Xenopus* XFD-10, XFD-2/2′ and rat HFH-5, XFD-5 and mouse fkh-4, as well as XFD-6, human H-8 and rat HFH-2 are orthologues. Additional sequence information for these proteins as well as the discovery of further genes from different organisms will be a pre-requisite to substantiate these notions on interspecies homologues.

In addition to its DNA binding property the winged helix domain of HNF-3β includes the nuclear localisation signal (Qian and Costa, 1995). It consists of H1 and the basic residues at the carboxyl-terminus (underlined in Table 1). Although both elements are fairly well conserved in all members of the winged helix family, the general significance of this complex element for nuclear transport remains to be shown.

4. Genomic organization and chromosomal locations

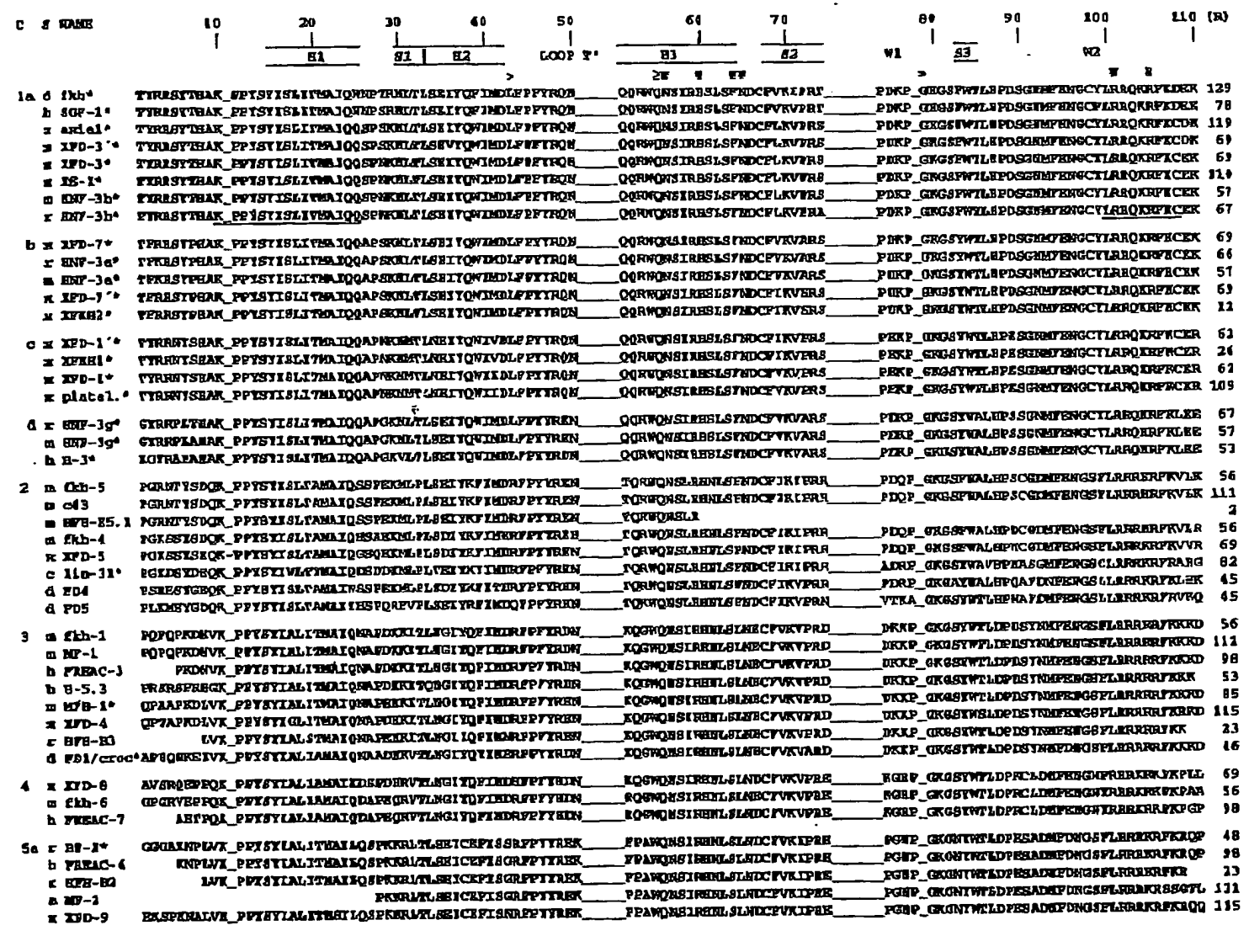
The fork head domain of winged helix proteins is often encoded by a single exon (e.g. *Drosophila* fkh, the HNF-3 family in rodents and most of the hitherto known XFD genes). However, the exons coding for the winged helix domains of rat HFH-3 and HFH-5 (Clevidence et al., 1993) are interrupted by an intron at a position corresponding to the W1 region. The genes for amphibian XFD-2/2′, XFD-10 (Lef et al., 1994; Scheucher et al., 1995) and the human freac-6 (Pierrou et al., 1994) contain an intron at the same position. Thus, all genes aligned to class 6 share this common feature. Other intron positions found are close to the carboxyl-terminal end of H2 in the *Drosophila* FD2 and FD3 genes (Häcker et al., 1992) or within the H3 motif of the rat HFH-4 gene. Interestingly, the latter position coincides with a locus where chromosomal rearrangements in alveolar rhabdomyosarcoma (Shapiro et al., 1993; Galili et al., 1993) and in acute lymphocytic leukaemia (Parry et al., 1994) had occurred.

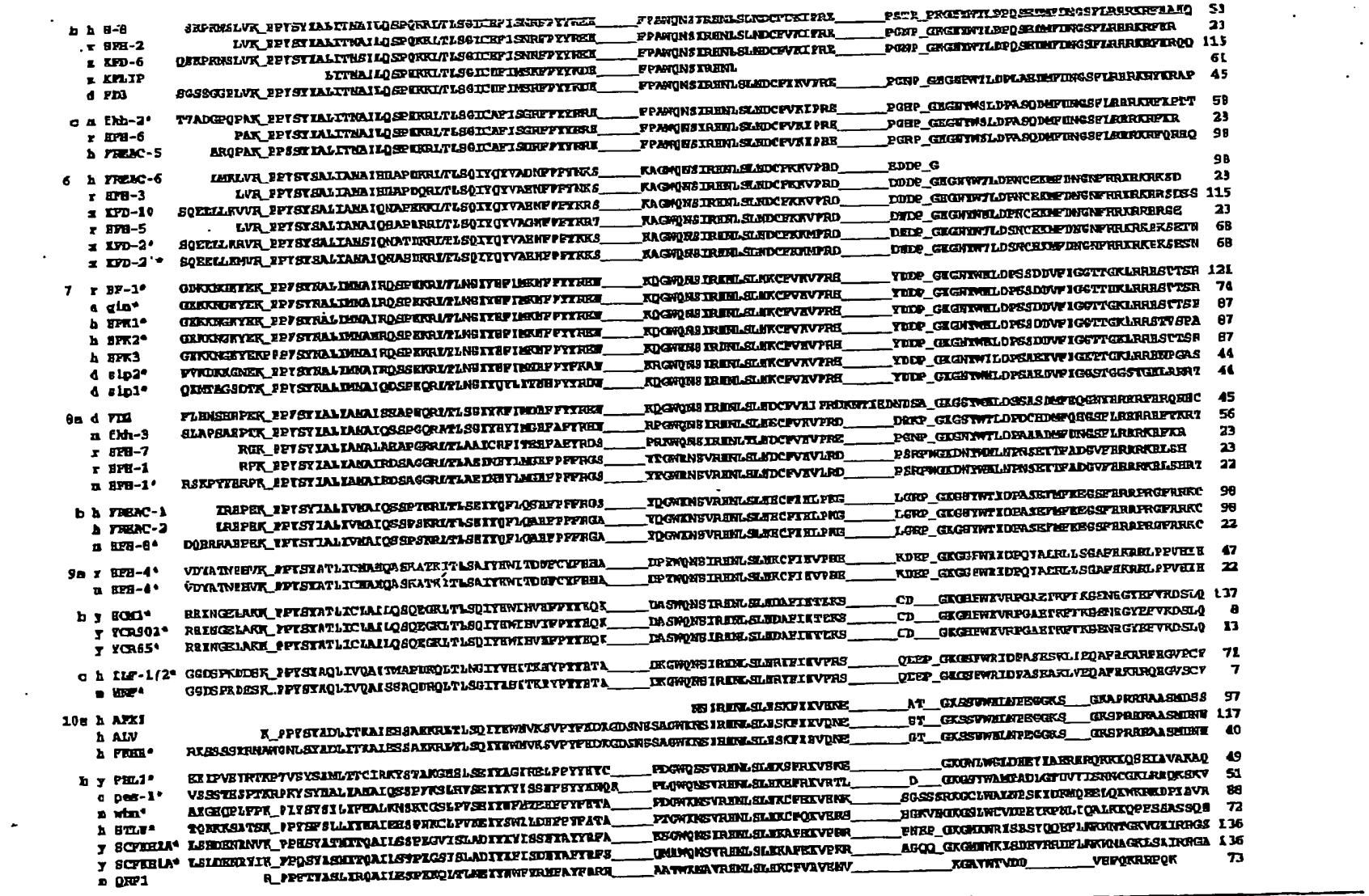
Some of the fork head genes, such as HNF-3β (Lai et al., 1990), BF-1 (Hatini et al., 1994) and PES-1 (Hope, 1994) give rise to multiple mRNAs. This is obviously due to variations within the transcription start site or to differential splicing of primary transcripts (Lai et al., 1993).

With respect to their chromosomal location the *Drosophila* genes FD1–3 map to separate regions of the genome (Häcker et al., 1992). The three related HNF-3 genes are spread over three different mouse chromosomes (HNF-3α: c12; HNF-3β: c2; HNF-3γ: c7) (Avraham et al., 1992). This suggests that the HNF-3 family has become widely dispersed during evolution and implies the necessity for independent activation of each member of this family. Mouse fkh-2 is located on chromosome 19 (Kaestner et al., 1995) and whn on chromosome 11 (Nehls et al., 1994). Similarly, the locations of various HFH genes (Clevidence et al., 1993, 1994) have been

Table 1

Alignment of fork head/HNF-3 domains of proteins from different organisms





Sequences were checked for last updates in DNA data libraries and categorized into ten different classes (C 1–10). In some cases, they were subdivided into sub-classes (a–d). While amino acids found to be most frequent for individual positions are shown in black, divergent residues are shown in class-specific colours. The following sequences are identical: XFD-3 = Xβ-1; XFKH2 = XFD-7′; XFKH1 = XFD-1′; XFD-1 = pintallavis; fkh-5 = o43 = HFH E5.1; fkh-1 = MF-1; BF-2 = HFH-B2; HCM1 = YCR902 = YCR65. ALV should be identical with FKHR despite the discrepancies reported for the amino-terminus. S denotes the source of sequences. a, avian; b, *Bombyx mori*; c, *Caenorhabditis elegans*; d, *Drosophila*; h, human; m, mouse; r, rat; x, *Xenopus*; y, yeast; z, zebrafish. (R) denotes references. Structural elements of the domain and positions (#) of direct base contacts (Clark et al., 1993) are shown on top. Intron positions for FD2 and FD3 are denoted by (>), for HFH-4 by (≥) and for all members of class 6 by (»). The nuclear localisation signal in HNF-3β is underlined.

determined on seven different chromosomes. BF-1 and BF-2 genes (Tao and Lai, 1992; Hatini et al., 1994) are both located on chromosome 12 but at a distance of 5.3 ± 2 cM apart from each other (Avraham et al., 1995). However, the two genes coding for the human brain factors HFK-1 and HFK-2 are found as a cluster on chromosome 14q (Wiese et al., 1995) and the *Drosophila* genes sloppy paired 1 and 2 map within some 10 kb of genomic DNA (Häcker et al., 1992). Thus, it seems that clustering of winged helix genes can only be expected for some very closely related genes, whereas the vast majority of fork head genes are located at different chromosomal loci.

5. Fork head domain proteins in *Drosophila*, *Caenorhabditis elegans* and yeast

Within the formal system of regulatory genes in *Drosophila*, the fkh protein belongs to the class of genes required for terminal pattern formation in the embryo. fkh mutants show defects in endodermal and ectodermal components of the gut, the nervous system and the salivary glands (Weigel et al., 1989). Thus, fkh is required for proper formation of the organs and tissues where the gene is expressed. Its in vivo regulation depends on redundant upstream and downstream cis-acting elements. The upstream elements extend into the transcribed regions of another gene (Weigel et al., 1990b). Genetic analysis revealed that the expression of fkh depends on the activated signal transduction pathway in response to the maternal gene torso, which is mediated by the activity of the two gap genes tailless and huckebein (Weigel et al., 1990a). Other genes, such as trithorax, which controls many genes in ANT and BX-C region (Kuzin et al., 1994) or sex combs reduced (Scr), a homologue of the rodent Hox-1.3 gene, are involved in the initial regulation of fork head expression at blastoderm (Zhao et al., 1993). During later development, fkh directs the expression of Krüppel in the posterior terminal region and, thereby, contributes to the formation of the malpighian tubules (Gaul and Weigel, 1990; Hoch et al., 1994). In addition, fork head may also be involved in the process of cell death (apoptosis) (Tepass et al., 1994).

In addition to fork head, seven fork head domain-encoding genes (FD1–7) were found to be expressed in the *Drosophila* embryo (Häcker et al., 1992). FD1 is detected in the early blastoderm corresponding to positions of precursor cells for anterior and posterior gut derivatives. During gastrulation, transcripts are present in a subset of cells within the central nervous system (CNS). FD1 turned out to be identical to the gene crocodile (croc) (Häcker et al., 1995). The mutant phenotype indicates that croc functions as an early patterning gene in the anteriormost blastoderm head segment anlage and establishes a specific head skeletal structure that derives from the non-adjacent intercalary segment at a later stage of embryogenesis. Genetic analysis revealed that croc activity in the head anlage is controlled by the anterior, the dorsoventral and terminal maternal organizer systems.

FD2 is found in a small number of cells in blastoderm; these cells invaginate during gastrulation (Häcker et al., 1992). Within the extended germ band stage ten segmental cell clusters with FD2 activity are visible; these cells represent most likely mesodermal tissue. FD3 transcripts are initially present in the fully extended germ band stage. Signals appear in neuroblasts which derive from the procephalic neurogenic region and in a subset of cells giving rise to the ventral nervous system. During later stages, distinct patterns evolve within the ventral nervous system, a subset of ventral midline cells and most prominently in a set of cells in the thoracic region. FD4 and FD5 transcripts first appear at fully extended germ band stage. These two genes are most prominent by the location of their transcripts within 14 symmetrical pairs of neuroblasts along the longitudinal axis of the embryo. The function of these genes is not known. FD6 and FD7 turned out to be identical to the segmentation genes sloppy paired (slp) 1 and slp2 (Grossniklaus et al., 1992). slp1 and slp2 are functionally redundant and biochemically equivalent (Cadigan et al., 1994a). The slp locus is required for the establishment of the metameric body plan and for the maintenance of anterior-posterior polarity of the parasegments, thereby involving segment polarity genes like wingless and engrailed (Cadigan et al., 1994b).

Fork head proteins are not only involved in morphogenesis but also in the process of cell specification. In *C. elegans*, lin-31 acts downstream of the *Caenorhabditis* ras-gene homologue let-60 in the regulation of the fate of vulval precursor cells (Miller et al., 1993) and PES-1 is also involved in cell signalling processes (Hope, 1994). The fork head proteins HCM1 (identical with YCR65 and YCR902: Benit et al., 1992; Bork et al., 1992) and FHL1 have been found in yeast. The HCM1 gene being localized on the recently sequenced chromosome three acts as a dosage-dependent suppressor of a temperature-sensitive calmodulin mutant (Zhu et al., 1993). Its action in this process is rather indirect and is not due to binding to the calcium carrier. FHL1 functions as a suppressor of the yeast RNA polymerase III (Hermann-Le Denmat et al., 1994). The gene is localized on chromosome 16 and its null mutation causes severe reductions of the cellular proliferation rate and the contents of rRNA. FHL1 is probably involved in processing of rRNA transcripts.

6. Vertebrates

The expression pattern of the HNF-3 family has been studied first in the adult. HNF-3α and HNF-3β are mainly found in liver and, to a lesser extent, in lung, intestine and stomach, whereas HNF-3γ is additionally present in heart, adipose tissue and testis, but not in lung (Lai et al., 1990, 1991; Kaestner et al., 1994). However, beyond its role as a liver-specific transcription factor the HNF-3 proteins do

exert a profound role at all stages of embryogenesis (Ang et al., 1993; Monaghan et al., 1993; Sasaki and Hogan, 1993). Mouse HNF-3α, β and γ are expressed sequentially during development of the definitive endoderm germ layer. HNF-3β is expressed at 6.5 dpc in the node at the anterior end of the primitive streak in all three germ layers. Within the next 24 h the streak elongates coupled with HNF-3β expression in the most distal point of the egg cylinder. The node gives rise to notochord (e.g. axial structures) and gut endoderm, both of which express HNF-3β. Finally, transcripts are localized in the floor plate and ventral midline cells of the CNS (McMahon, 1994). Within the course of body axis formation HNF-3β seems to induce the expression of sonic hedgehog in notochord, floor plate and forelimb buds. In turn, sonic hedgehog seems to maintain the expression of HNF-3β (Echelard et al., 1993; Riddle et al., 1993; Sasaki and Hogan, 1994).

The transcription of HNF-3α is initiated later than HNF-3β in midline endodermal cells in the region of the invaginating foregut. Subsequently, it is apparent in notochord, ventral neural tube and gut endoderm, similarly to HNF-3β. HNF-3γ then appears upon hindgut differentiation. The different boundaries of HNF-3α, β and γ during endoderm development suggest that they have a crucial function in endoderm regionalization. Additionally, the HNF-3 family is found within the developing mesoderm and structures derived thereof. HNF-3α and β are present in the cells of the floor plate in ventral neural epithelium and in the chordamesoderm (notochord). At later stages the HNF-3 transcripts are concentrated to an increasing extent in endodermal tissues, such as the liver.

The *Xenopus* organizer and its equivalent region in mouse (node) or chick (Hensens node) emits a signal for the dorso-ventral patterning of forming mesoderm (somites) and a planar signal initiating the patterning of the anterior-posterior axis in the neural plate. In later stages of embryogenesis signals from the organizer specify cells that give rise to specific midline structures, such as the notochord which has also the capacity in forming dorso-ventral polarity of the spinal chord. The location of winged-helix proteins within the *Xenopus* organizer, such as XFD-1/1′ (pintallavis/XFKH1) (Dirksen and Jamrich, 1992; Knöchel et al., 1992; Ruiz i Altaba and Jessell, 1992), HNF-3β in mouse node (Monaghan et al., 1993) or axial (Strähle et al., 1993) in the zebrafish organizer suggests that they may be crucial for both anterior-posterior and dorso-ventral axis formation, respectively. But while ectopic expression of the organizer-enriched homeobox protein goosecoid causes a second axis (Cho et al., 1991), such a dramatic effect is not described for the organizer-enriched fork head proteins. However, the engrailed promoter-directed expression of HNF-3β in mouse midhindbrain induces ventral structures in dorsal brain regions (Sasaki and Hogan, 1994) and, correspondingly, by ectopic expression of frog pintallavis the expression of F-spondin, a floor-plate-specific marker, is observed in dorsal regions (roof plate) of the hindbrain (Ruiz i Altaba et al., 1993a). Moreover, partial or complete HNF-3β knock-outs (Weinstein et al., 1994; Ang and Rossant, 1994) have profound consequences in morphogenesis. Homozygous mice show the first abnormalities at 6.5 dpc and die at 9.5 dpc. Gastrulation is initiated but the egg cylinder is deformed and the anterior-most cells of the primitive streak are dispersed. The available data suggest that a proper node, notochord and its extension towards the brain (the head process) will never form, while other mesodermal-derived tissues (somites) are generated. This is probably due to some remnants of node structures, since node-derived gut endoderm is formed (Ang and Rossant, 1994; Weinstein et al., 1994). The development of gut seems to be normal apart from failures in foregut closure. Therefore, HNF-3β is probably not required for initial gut formation, even if it is later expressed in this tissue. Since HNF-3α precedes the expression of HNF-3β in gut endoderm, one might speculate that HNF-3α is the candidate gene in gut formation. Surprisingly, the development of the neural plate can also be followed in the mutant in the absence of notochord. This might be due to the presence of planar signaling within the anterio-posterior axis in the absence of vertical signaling (Ruiz i Altaba and Jessell, 1992; McMahon, 1994).

The *Xenopus* fork head proteins XFD-1/1′ (pintallavis, XFKH1) (Dirksen and Jamrich, 1992; Knöchel et al., 1992; Ruiz i Altaba and Jessell, 1992) are activated as a direct zygotic response to dorsal mesoderm induction in the dorsal blastopore lip (Spemann organizer). Transcripts are later detected in notochord and, probably as a direct response to signals emitted from the notochord, in the midline neural plate cells (neural floor plate). The potential role of XFD-1 in mesoderm formation is supported by its induction with activin A in the absence of de novo protein synthesis. Like the dorsal lip-specific homeobox gene goosecoid (Painsod et al., 1994), XFD-1/1′ is down-regulated by bone morphogenetic proteins (Clement et al., 1995). Overexpression of pintallavis causes perturbation in neural axis formation. It suppresses the development of anterior and dorsal neural cells and leads to an expansion of the posterior neural tube (Ruiz i Altaba and Jessell, 1992). It has to be noted that *Xenopus* XFD-1 with respect to its primary sequence and lack of late expression is not the direct equivalent of mammalian HNF-3β. A *Xenopus* homologue of rodent HNF-3β, Xβ-1 (being identical with XFD-3) (Knöchel et al., 1992; Ruiz i Altaba et al., 1993b), has been isolated. Its late expression is similar to HNF-3β but it is not transcribed in notochord. Vice versa, a mammalian XFD-1 homologue could not be identified. Since XFD-1 and XFD-3 share a high degree of sequence similarity within their fork head domains and their combined expression patterns correspond to that of mammalian HNF-3β, it was suggested that mouse HNF-

3β functionally accounts for the combined action of the two frog proteins (Ruiz i Altaba et al., 1993b).

The zebrafish fork head gene axial is detected just before gastrulation in the zebrafish organizer (Strähle et al., 1993). Expression is found in involuting cells comprising the mesendoderm of the developing axis. At the end of gastrulation transcripts are found in mesodermal and adjacent ventral neural plate cells along the body axis. Similar to XFD-1 in frog and HNF-3β in mouse, axial seems to be involved in the development of dorsal mesoderm and axis formation. The activation of axial expression by activin further argues for a functional similarity of these genes. The cyclops mutation in fish causes, by lack of the neural floor plate, a deficiency in brain development. Since these anomalies correlate with domains of axial transcription, it is likely that axial responds to cyclops gene activity. Axial acts in concert with eph, pax and wnt activities in brain specification in such a manner that 'scaffolds' are established for proper axon guidance in the process of neuron differentiation (Macdonald et al., 1994).

Another six HNF-3-related fork head genes, fkh-1 to 6, were reported in mouse (Kaestner et al., 1993). Transcripts are present from 9.5 dpo and are later found in derivatives of all three germ layers. A more recent analysis revealed that fkh-2 is expressed in headfold stage embryos in the notochord, the anterior neuroectoderm and a few cells of the definitive endoderm (Kaestner et al., 1995). This expression becomes restricted to anteriormost portions of invaginating foregut and the developing midbrain. Later, transcripts are restricted to midbrain and become progressively localized to red nuclei. Similarly, another set of HNF-3 related proteins from rat, HFH-1 to 7, has been described (Clevidence et al., 1993). Transcripts are found in adult liver, brain, kidney, lung and intestine. Mouse HFH-E5.1, being identical to fkh-5, is transiently expressed during gastrulation in posterior ectoderm and mesoderm, but is not present in endoderm (Ang et al., 1993). MFH-1 is a mouse fork head protein which is predominantly found in mesenchyme from 9.5 dpo onwards, first in non-notochordal mesoderm (somites), later in areas of mesenchymal condensation (trunk, head, limbs) and in bone primordia: it is not found in CNS, heart, liver, lung and gut (Miura et al., 1993). Postnatally (day 3), the protein is detected in kidney and skin. Mouse MF-1 (identical to fkh-1) and MF-2 show predominant expression in mesoderm (Sasaki and Hogan, 1993). MF-1 is localized in non-notochordal mesoderm and head mesenchyme, whereas MF-2 is present in head mesenchyme and sclerotomes of the somites. In mouse gastrulation, the restricted expression patterns of HNF-3β, MF-1 and MF-2 in the node correlate with cell populations being determined for different fates.

The fork head protein BF-1 was detected in mouse CNS (Tao and Lai, 1992). Its expression is unique to the telencephalon of the developing brain; transcripts are already detectable at 10 dpc. They are restricted to an area of the neural tube which later develops into telencephalic vesicles (Tao and Lai, 1992). The crucial role in mouse forebrain development was underlined by BF-1 knockouts (Xuan et al., 1995). Whereas the heterozygous mutant develops normally, the homozygous mutant dies at birth and shows strongly reduced cerebral hemispheres. Many cells of the neuroepithel show premature differentiation. Thus, BF-1 probably controls telencephalon morphology by regulating the rate of cell proliferation and the proper timing of the differentiation of the cerebral hemispheres. Expression of BF-2 (Hatini et al., 1994) is restricted to the rostral diencephalic neuroepithel. Rostrally, the domain of BF-2 is adjacent to BF-1 and the boundary extends laterally to divide optic stalk and retina into anterior and posterior areas. The expression domains of the two factors thus define a boundary in the developing forebrain which plays a role in the subdivision of the forebrain and establishes positional information in the retinal neuroepithel. The avian sarcoma virus oncogene, qin, seems to be the avian homologue to BF-1 (Li and Vogt, 1993) and the human homologue to BF-1 has been described as HFK1 (Murphy et al., 1994; Wiese et al., 1995). Its expression in human foetal brain is restricted to the neuronal cells in telencephalon, with high abundance in the developing dentate gyrus and hippocampus. Two other human brain fork head proteins, HFK2 and 3, are related but different from HFK1 (Murphy et al., 1994).

Including the already mentioned XFD-1, a total of ten different *Xenopus* fork head domain-related proteins, XFD-1 to 10, has been described (Knöchel et al., 1992; Lef et al., 1994, 1996; Scheucher et al., 1995). XFD-2 is activated at the early blastula stage within the animal half of the embryo (Lef et al., 1994). At late blastula, transcripts are concentrated in the marginal zone. The gene is expressed in all parts of forming mesoderm. Transcripts are later found in somites, notochord, neural floor plate, brain and spinal chord. The early expression of XFD-2 is complemented by the early pattern of the *Xenopus* homologue to mammalian HNF-3α, XFKH2 (Bolce et al., 1993) being identical to XFD-7′ (Lef et al., 1996). Transcripts are found in the vegetal hemisphere and within the marginal zone, but not in the animal pole region. Equal amounts of transcripts are subsequently localized in the dorsal and ventral halves of gastrula and neurula stage embryos. The dorsal component is concentrated in notochord (chordamesoderm) at neurula stage. At tailbud stage, XFKH2 transcripts are present in foregut, brain and floor plate. The early transcription of this gene can be induced by activin A in the absence of de novo protein synthesis, suggesting that XFKH2 acts as another early response gene in the process of mesoderm induction (Bolce et al., 1993). XFLIP is a *Xenopus* fork head-related gene which shares a high degree of sequence homology with *Drosophila* FD3. It is transcribed in the dor-

sal blastopore lip and is induced in animal caps by activin and by bFGF (King and Moore, 1994).

XFD-4 and XFD-6 are activated already in early gastrula stage embryos, whereas XFD-9 and XFD-10 are later transcribed, reaching their highest activity at tailbud stage (Lef et al., 1994). XFD-4 transcripts are first detected in somitogenic mesoderm, at tailbud stage in pronephros and at tadpole stage in heart and in the tip of the tail. XFD-6 expression is most prominent in migrating neural crest cells originating from mesencephalon and rhombencephalon. XFD-9 transcripts are detected at somite segregation stages in the region of rotating somites and within the posterior part of the pronephric duct. XFD-10 is activated in neuroectoderm and somitogenic mesoderm, but, at tailbud stage, transcripts are restricted to distinct cell lines at hyoid and anterior branchial arches, respectively (Scheucher et al., 1995).

The regulation of the myoglobin gene by the fork head protein MNF in the course of mammalian myogenesis is another example for a winged helix protein in cellular specification (Bassel-Duby et al., 1994). Similarly, cell lineage within the human hematopoietic system seems to be affected by fork head proteins H3 and H5–3 (Hromas et al., 1993; Hromas et al., 1994). The mouse whn gene accounts for the nude mice phenotype showing disrupted thymus development and loss of hairs (Nehls et al., 1994). A checklist for all of these fork head/HNF-3 proteins is shown in Table 2.

7. Tumorigenesis

The first oncogene which was reported to encode a protein of winged helix structure was the avian sarcoma virus 31 gene qin (Li and Vogt, 1993). qin determines the transforming activity of the retrovirus and displays a particular homology to BF-1.

A causative role of winged helix proteins in chromosomal aberrations leading to tumorigenesis has been found in acute lymphocytic leukaemia and paediatric alveolar rhabdomyosarcoma. One type of acute lymphocytic leukaemia is caused by a translocation between the X chromosome and chromosome 11 [t(X,11)]. This translocation involves the fusion of the general transcription factor HTRX1, a Zn-finger protein, on chromosome 11 to a fork head gene, AFX1, on the X chromosome (Parry et al., 1994). In alveolar rhabdomyosarcoma, a translocation fuses the pax3 gene on chromosome 2 with the fork head gene ALV on chromosome 13 [t(2,13)]. The pax3 gene is normally a potent activator found to be expressed in lateral dermomyotome which later forms limb muscle progenitors. The paired box and the homeodomain of pax3 are fused to helix 3 of the fork head domain of ALV (Shapiro et al., 1993) or FKHR (Galili et al., 1993) in such a manner that the pax transactivation domain is replaced by the remaining part of the fork head domain. This fusion product was shown by in vitro as well as in vivo experiments to be a stronger activator than the pax3 protein (Fredericks et al., 1995; Sublett et al., 1995). As pax3 in myoblast cell cultures is able to inhibit myogenic differentiation, the pax-fork head fusion in the sarcoma suppresses terminal differentiation of migrating limb myoblasts (Epstein et al., 1995). Therefore, these fusions contribute to the malign phenotype by preventing terminal differentiation and subverting the process of normal cell development (Galili et al., 1993; Shapiro et al., 1993; Epstein et al., 1995). RT-PCR-based diagnoses to assess this chromosomal translocation are established (Downing et al., 1993).

A region on human chromosome 17q25, a site of translocation involved in acute myelogous leukaemia, encodes the interleukin enhancer binding factor (ILF), a fork head protein involved in HIV infection (Li et al., 1991, 1992a). HIV activation is regulated by factors induced in the course of T-cell activation such as NFκB or the nuclear factor of activated lymphocytes (NFAT) which binds to the interleukin-2 promoter and the HIV long terminal repeat. ILF was found as an additional factor in the regulation of this locus and is expressed constitutively in lymphoid and non-lymphoid tissues. Similarly, the long terminal repeat of the human T-cell leukaemia virus is important in the regulation of gene expression by transcription factors of the ets family (Macleod et al., 1992). Human T-cell leukaemia virus enhancer factor (HTLF), a fork head protein located on chromosome 2, is involved in this process (Li et al., 1992b). As already noted, there is an interesting structural homology between the ets transcription factors and the winged helix family.

8. Recognition sites

As could be expected from their similar structure, a multitude of known DNA target sites for different winged helix proteins share a common recognition motif. In vitro target sites for rodent HNF-3β, HFH-1, HFH-2 (Overdier et al., 1994), human fork head-related proteins FREAC-1 to 7 (Pierrou et al., 1994), *Drosophila* crocodile (Häcker et al., 1995) and *Xenopus* fork head-related proteins, XFD-1, -2 and -3 (Kaufmann et al., 1995) were determined by means of a PCR-assisted binding site selection procedure employing degenerated deoxyoligonucleotides. Roux et al. (1995) extracted a consensus binding site from reported HNF-3 binding sites. All these studies have a 7 bp recognition core motif 5′ [(G/A) (T/C) (C/A) A A (C/T) A] 3′ in common, whereas sequences flanking either side do not share any obvious similarity. The core region on its own is necessary but not sufficient for protein recognition in vitro. The terminal residues of the core motif and the flanking sequences contribute to binding specificity (Overdier et al., 1994; Pierrou et al., 1994; Kaufmann et al., 1995; Roux et al; 1995). However, recognition does not seem to be due to a simple sequence

Table 2
Checklist of fork head/HNF-3 genes

Species	Gene	Embryonic (EE) or adult expression (AE), knock out phenotype (KO), comments (C)	References
Yeast	HCM1	C: HCM1 suppressor of calmodulin mutant	[137]
	YCR902	C: chromosome III	[8]
	YCR65w	C: chromosome III	[13]
Yeast	FHL1	C: suppressor of conditional mutation in RNA polymerase III	[49]
C. elegans	lin-31	C: involved in vulval development, acts downstream of ras homolog let-60	[82]
C. elegans	PES-1	EE: early embryogenesis, anterior ventral surface of embryo C: cell decision in early embryogenesis	[51]
Drosophila	FKH	EE: early blastoderm, terminal regions, gut precursor cells, at later stages in nervous system, salivary glands C: mutants gut structures replaced by ectopic head structures	[129]
Drosophila	FD1–7	EE: FD1,2 first in early blastoderm, FD1 gut precursor cells, FD2 extended germ band 10 segmental mesodermal cell clusters, FD3,4,5 first in extended germ band stage, FD3 neuroblasts, ventral nerve system, cell cluster in thoracic region, FD4,5 in 14 symmetrical pairs of neuroblasts C: FD1 = croc; FD6,7 = slp1,2	[45]
Drosophila	slp1, 2	EE: very similar expression patterns of slp1, 2, end of cellular blastoderm segment polarity gene-like expression C: involved in maintenance of segment polarity, biochemically redundant	[44] [15,16]
Drosophila	croc	EE: early patterning gene in anteriormost blastoderm head anlage, establishment of specific head skeletal structure C: controlled by anterior, terminal and dorsoventral organizer systems	[46]
B. mori	SGF-1	EE: at all larval stages in middle silk glands C: regulator of sericin-1 gene in silk glands	[78]
Zebrafish	Axial	EE: present in all germ layers, axial mesendoderm and ventral nervous system C: Activin inducible expression, acts downstream of the cyclops locus	[119]
Xenopus	XFKH1	EE: organizer, posterior notochord, neural floor plate C: activin-inducible expression, identical to XFD-1′	[26]
Xenopus	XFD-1/1′	C: pseudoalleles; see Pintallavis/XFKH1 downregulation by BMP	[62] [21]
Xenopus	Pintallavis	EE: see XFKH1 C: identical to XFD-1, overexpression leads to expansion of posterior neural tube and to ectopic expression of F-spondin	[109] [108]
Xenopus	Xβ-1	EE: dorsal lip, floor plate but not in notochord, pharyngeal endoderm C: identical to XFD-3, frog homolog to HNF-3β, in contrast to mammalian HNF-3β onset of expression in neural tissue after closure of neural tube	[110]
Xenopus	XFKH2	EE: at blastula in vegetale pole, later notochord, at tailbud stage in brain, foregut, floor plate C: identical with XFD-7′, frog homolog to HNF-3α, activin-inducible expression	[12]
Xenopus	XFD-2/2′	EE: at early blastula in animal hemisphere, at late blastula in marginal zone, later in chordamesoderm, neurocctoderm	[68]
Xenopus	XFD-3	EE: dorsal lip C: activin and bFGF inducible expression, related to *Drosophila* FD3	[61]
Xenopus	XFD-4, 6, 9, 10	EE: XFD-4 in somitogenic mesoderm, later in pronephros, XFD-6 in migrating neural crest cells, XFD-9 pronephric duct, rotating somites, XFD-10 in hyoid and anterior branchial arches	[115]
Xenopus	XFD-3/3′	C: pseudoalleles, frog homologues to mammalian HNF-3β, identical to Xβ-1	[62,69]
Xenopus	XFD-7/7′	C: pseudoalleles, frog homologues to mammalian HNF-3α, identical to XFKH2	[69]
Chick	qin	C: avian sarcoma virus oncogene, homolog to BF-1 and HFK1	[74]
Rat	HNF-3α, β, γ	AE: liver, small amount in intestine, absent from brain, spleen and kidney, HNF-3α, β in lung, HNF-3γ in testis	[66,67]
Rat	HFH-1–7	AE: brain, heart, lung, liver, intestine, kidney, not in spleen	[23]
Rat	HFH-4	EE: proximal pulmonary epithel AE: oviduct, testis C: rat homolog to mouse HFH-4, involved in lung development and spermatogenesis	[47]
Rat	BF-1	EE: developing neural tube at rostral end giving rise to telencephalon KO: reduced cerebral hemispheres C: involved in telencephalon development, homolog to qin and HFK1	[121] [134]
Rat	BF-2	EE: in CNS restricted to rostral diencephalic neuroepithelium C: in concert with BF-1 involved in forebrain regionalization	[48]
Mouse	fkh-1 to 6	EE: present during embryogenesis	[56]

		AE: fkh-1 brain, heart, kidney, fat; fkh-2 lung, spleen: fkh-3 lung, gonads; fkh-4, 5 brain, thymus; fkh-6 lung, kidney, stomach, intestine	
Mouse	fkh-2	EE: anteriormost portions of the invaginating foregut and developing midbrain, later restricted to the midbrain and red nuclei C: relates to HFH-6	[58]
Mouse	MFH-1	EE: non-notochordal mesoderm, later in areas of mesenchymal condensation in trunk, head and limbs, before bone formation	[85]
Mouse	QRF-1	C: glutamine-rich factor, involved in terminal differentiation of B-cells	[73]
Mouse	MF-1, 2	EE: MF-1 in non-notochordal mesoderm, neural crest derived head mesenchyme; MF-2 in sclerotomes of somites and head mesenchyme	[111]
Mouse	MNF	AE: heart, skeletal muscle C: regulation of myoglobin gene within myogenic differentiation	[7]
Mouse	HFH-1, 4, 8	EE: HFH-1 differentially expressed in lung, HFH-4 in Clara cells, HFH-8 in type II pneumocyte cells C: HFH-8 in concert with HNF-3α involved in surfactant protein B regulation	[22]
Mouse	whn	AE: thymus, skin C: mutation impairs thymus development and hair growth (nude mouse)	[88]
Mouse	HNF-3α, β, γ	EE: HNF-3α, β in endoderm-derived tissues, HNF-3γ additionally in ovary, testis, heart, but not lung KO: HNF-3β abnormal node, no notochord, other mesodermal cell types are formed C: mouse equivalents to rat HNF-3 family	[2,57,86,111,112] [1,131]
Mouse	HFH-E5.1	EE: expressed in posterior ectoderm and mesoderm at primitive streak stage	[2]
Human	ILF-1, 2	AE: constitutively in lymphoid and non-lymphoid tissues C: binds to HIV-long terminal repeat, IL-2 promoter	[70,71]
Human	HTLP	C: binds to human T cell leukaemia virus long terminal repeat	[72]
Human	FKHR	C: paediatric alveolar rhabdomyosarcoma, fusion with PAX3 due to translocation [t(2,13)], identical to ALV	[40]
Human	ALV	see FKHR	[117]
Human	AFX1	AE: in cell line from acute lymphocytic leukaemia C: fusion with HTRX1 due to translocation [t(X,11)]	[97]
Human	H3, H8, H5–3	AE: H8 ubiquitously expressed, H3 and H5–3 in human hematopoietic cell lineage	[52,53]
Human	FREAC-1 to 7	EE: FREAC-1, -2 lung; FREAC-3, -5 brain, liver, kidney; FREAC-4 kidney; FREAC-6 kidney AE: FREAC-1, 2 placenta, lung; FREAC-3 ubiquitous; FREAC-4 monocyte cell lineage; FREAC-5 muscle tissue: FREAC-6 kidney	[98]
Human	HFK1, 2, 3 HBF-1, 2, 3	EE: HBF1 in telencephalon neural cells, developing dentate gyrus and hippocampus C: HFK1 homologous to qin and BF-1	[87]
Human	HBF-2	EE: brain AE: brain, muscle, testis C: clustered with HBF-1 on chromosome 14	[132]

pattern. The DNA bound to the FREAC proteins (Pierrou et al., 1994) is induced to a strong bend of 90° upon protein binding, suggesting that such a steric change might influence binding properties. Possibly, recognition specificity is conferred by charge patterns or more elaborate structures within the flanking regions.

Correspondingly, the binding domain of the protein contains a region which confers binding specificity. Domain-swapping and mutagenesis experiments revealed that binding specificity is dependent on a stretch of 20 amino acid residues between the carboxyl-terminal end of H2 and the amino-terminus of H3. Surprisingly, according to the X-ray analysis of HNF-3γ this region is not in contact with DNA (Overdier et al., 1994).

Due to the crystallographic analysis of the complex of the HNF-3γ fork head domain with 11 residues from the TTR promoter (plus two artificial C residues) the centre of the protected region encompasses a 7 bp sequence *T* A A G T C A (italic: in contact with minor groove; underlined: in contact with major groove) (Clark et al., 1993). The alignment of this sequence with the consensus core sequences derived from in vitro binding site selection experiments (Pierrou et al., 1994; Kaufmann et al., 1995; Roux et al; 1995) indicates that the latter sequences may encompass more binding residues than those employed in the X-ray study with the 11 bp TTR promoter motif. This assumption is substantiated by the fact that the site selected for the FREAC proteins is bent by 90° (Pierrou et al., 1994) and not by 13° as found in the diffraction study with the TTR promoter motif, by the finding of a footprint for the *Drosophila* and *Xenopus* fork head domains which extends beyond the corresponding region in the 'crystal' TTR sequence (Kaufmann et al., 1994, 1995) and the usage of at least one extra residue at the 3′ end of the TTR target which is not part of the natural sequence. Another detailed comparative study on the X-ray data led to a similar conclusion (Roux et al., 1995). Thus, structural investigations on fork head domains including more extended recognition sites are necessary to obtain a complete resolution of protein/DNA interaction.

Besides the well known sites for HNF-3 within the TTR promoter, corresponding target sites have been

found in the 5′ region of the α-fetoprotein (Costa et al., 1989; Millonig et al., 1995) and the α-antitrypsin (elastase) promoter region (Costa et al., 1989). By comparison of known HNF-3 recognition sites to computer data banks, the number of potential HNF-3 binding sites has remarkably increased (Overdier et al., 1994; Roux et al., 1995). However, not all these putative recognition sites have yet been experimentally verified (e.g. Eguchi et al., 1991; Fahrner et al., 1993; Shaw et al., 1994, Tremp et al., 1995). Those targets which have been experimentally studied were mostly analysed by virtue of their binding to HNF-3. In many cases, the binding capacity of the DNA has been inferred from their sequence similarity to the TTR promoter and subsequent gel retardation assays including specific competition and 'super-shift' experiments employing HNF-3 antisera.

Only in a few cases has the molecular identity of the binding protein been directly verified by classical protein-chemical analysis, such by identification of HNF-3α and HNF-3β as binding factors in the rat pancreatic α-amylase promoter (Cockell et al., 1995) or by in vivo footprinting (Cardinaux et al., 1994) on the previously in vitro characterized A-activation protein binding site in the *Xenopus* A2 vitellogenin promoter (Drewes et al., 1991).

More HNF-3 recognition sites have been identified in viral genes, such as in the hepatitis B virus enhancer (Chen et al., 1994; Ori and Shaul, 1995), the promoter of the hepatitis B virus large surface antigen (Raney et al., 1995) and the duck hepatitis B virus enhancer (Liu et al., 1994). An autoregulatory binding site for HNF-3β has been located in the HNF-3β upstream region (Pani et al., 1992b). Many genes for key enzymes within metabolism contain HNF-3 binding sites, e.g. the promoter of the human lipoprotein lipase gene (Enerbäck et al., 1992), the 5′ region of the phosphoenolpyruvate carboxykinase gene (Ip et al., 1990; Angrand et al., 1994), the −2.5 and −11 kb enhancer of the tyrosine amino transferase promoter (Nitsch and Schütz, 1993; Nitsch et al., 1993a) and the aldolase proximal promoter (Raymondjean et al., 1991; Gregori et al., 1993, 1994). Other target sites include the promoter of the Clara cell secretory protein (Bingle and Gitlin, 1993; Sawaya et al., 1993; Sawaya and Luse, 1994; Bingle et al., 1995), the regulatory region of the lung surfactant protein B (Bohinski et al., 1994), the transferrin gene (Auge-Gouillou et al., 1993) and the liver-specific enhancer of serum albumin gene (Jackson et al., 1993).

Besides these HNF-3-type binding sites only a few natural target sites for other fork head proteins were identified. Amongst them is the binding of the *Bombyx mori* silk gland factor to the sericin-1 gene promoter (Mach et al., 1995) and the binding of *Drosophila* fork head to the Krüppel gene promoter (Kaufmann et al., 1994). Detected target sites share similarities to those of the HNF-3 family. However, binding of the myocytic nuclear factor to the CCAC element within the human myoglobin gene (Bassel-Duby et al., 1994), the interleukin enhancer binding factor to the HIV long terminal repeat and the interleukin-2 gene promoter (Li et al., 1991; Li et al., 1992a), and binding of the human T cell leukaemia virus enhancer factor HTLF to the human T cell leukaemia virus long terminal repeat (Li et al., 1992b) revealed completely different sites, as could be expected from their remote sequence relationship to the HNF-3 factors.

Since most of the newly found members of the winged helix gene family have been identified by virtue of their highly conserved DNA binding domain, the role of the flanking regions has not been studied in such detail. HNF-3β contains each two elements at its carboxyl-terminal and amino-terminal end which are involved in transcriptional activation (Pani et al., 1992a; Qian and Costa, 1995). The two regions at the carboxyl-terminus act position-independently, whereas the amino-terminal elements require the presence of the carboxyl-terminal domains. The transcriptional activity of the two amino-terminal regions is dependent on the integrity of a putative α-helical structure. These activation modules can in part also be found in other winged helix proteins, such as HNF-3γ (Lai et al., 1991), BF-1 (Tao and Lai, 1992), BF-2 (Hatini et al., 1994), XFD-1/1′ (Knöchel et al., 1992), fkh (Weigel et al., 1989), axial (Strähle et al., 1993), slp-1, slp-2 (Grossniklaus et al., 1992), lin-31 (Miller et al., 1993) and SGF-1 (Mach et al., 1995). The amino-terminus in HNF-3β contains a putative site for casein kinase, but phosphorylation does not seem to be essential for transcriptional activity (Qian and Costa, 1995). In contrast, the QRF1 factor has been found to be phosphorylated in vivo. This is the only known case where the activity of a fork head protein seems to be regulated by phosphorylation (Bassel-Duby et al., 1994).

9. Gene regulation by winged helix proteins in differentiated tissues

The different members of the HNF-3 family are involved in numerous regulatory circuits. In concert with HNF-4, HNF-3β determines the expression of HNF-1 in differentiated hepatocytes (Kuo et al., 1992). Binding of HNF-3 to cis-regulatory elements in the promoter of the lipoprotein lipase gene is associated with the formation of adipocytes from preadipocytes (Enerbäck et al., 1992). HNF-3β and γ cooperate with pancreas transcription factor 1 in exocrine pancreatic cells by binding to the 5′ region of the α-amylase 2 gene (Cockell et al., 1995). Similarly, phosphoenolpyruvate carboxykinase is activated by the action of HNF-3 on a −4.8 kb enhancer element (Ip et al., 1990). The eH site within the albumin enhancer is cooperatively activated by HNF-3α and NF-1, a phenomenon which is only observed within the natural context of the albumin enhancer (Jackson et al., 1993). Again, the most distal enhancer of the α-fetoprotein gene confers tissue specificity by HNF-3 binding (Millonig et

al., 1995). In liver-derived HepG2 cells, the HNF-3 site within the TTR promoter is partially overlapped by an AP-1 binding site, but these two proteins act in a non-cooperative fashion. During partial hepatectomy the levels of HNF-3α and TTR remain nearly unchanged. In case of acute phase liver (trauma, infection), HNF-3α and TTR are strongly down-regulated (Qian et al., 1995). The autoactivation of HNF-3β has already been mentioned. Interestingly, HNF-3α acts in hepatocytes as a strong anti-activator in the context of the aldolase promoter and antagonizes the role of D-box binding protein and HNF-1 (Gregori et al., 1993, 1994).

The development of the lung is an interesting example of the regulation of endoderm derived cells by HNF-3. The pulmonary epithel originates from an outpouching of embryonic gut endoderm into the surrounding foetal mesenchyme. The specification of this epithelial layer gives rise to various cell types (Bingle and Gitlin, 1993; Hackett et al., 1995) which are partially characterized by marker proteins, whose function is hitherto unknown. The airway cells in the bronchiolar epithel (Clara cells) express the CC10 gene product, whereas the alveolar epithel is characterized by expression of the surfactant protein C; surfactant proteins A and B are present in both cell populations (Sawaya et al., 1993). Whereas HNF-3α and β are co-expressed in liver, they are differentially expressed in lung. Together with HFH-4, HNF-3α is found in Clara cells, while HNF-3β is present in smooth muscle cells surrounding the bronchioles (Clevidence et al., 1994). HFH-8, another lung-specific factor, is found in type II pneumocytes. The localisation of surfactant protein B overlaps with the expression patterns of HFH-8 and HNF-3α. Indeed, it was demonstrated that the surfactant protein B expression is regulated by HNF-3α and HFH-8. This is the first example where cell type specificity is achieved by two members of the same family of transcription factors (Clevidence et al., 1994). The surfactant protein B promoter has additional target sites for the thyroid transcription factor 1 (TTF1). This finding, together with the fact that liver, lung and thyroid specification are determined by factors such as HNF-3α, β, γ, TTF1 and pax8, led to the hypothesis that organ-specific expression along the foregut axis is determined by combinations of common activators (Bohinski et al., 1994). The 5′ region of the CC10 gene contains within 70 bp binding sites for AP1, octamer factors and HNF-3 (Sawaya et al., 1993). Surprisingly, co-transfection experiments with CC10 promoter elements and HNF-3α and β in HeLa cells resulted in activation by HNF-3α and repression by HNF-3β. Thus, two members of the same family exert opposite effects (Sawaya and Luse, 1994). However, this result seems to be specific for HeLa cells. By similar co-transfection studies in NCI-H441 cells it was found that the two factors act in a cooperative manner (Bingle et al., 1995).

The tyrosine aminotransferase (TAT) gene in rodent liver is activated by glucocorticoids and glucagon and is repressed by insulin (Nitsch and Schütz, 1993; Nitsch et al., 1993a). Gene expression is mediated by three enhancer elements. The most distal enhancer at −11 kb responds to HNF-3α and γ, but not to HNF-3β. A −2.5 kb element binds to HNF-3α, β and γ. The combination of signal transduction and specific transcription factors ensures a proper developmental TAT activation. Within the −2.5 kb region the glucocorticoid receptor (GR) and HNF-3 response element overlap such that the level of HNF-3 influences the contribution of the GR binding sites (Roux et al., 1995). The presence of ets recognition sites could be responsible for non-steroidal extracellular stimuli on TAT gene expression (Espinas et al., 1994, 1995). Further, extinction of TAT gene activity in somatic cell hybrids (hepatoma/fibroblast) involves modification and loss of several transcriptional activators, such as HNF-4 and HNF-3α, β and γ (Nitsch et al., 1993b).

10. Conclusions and outlook

The analysis of the three-dimensional structure of the HNF-3γ fork head domain led to the discovery of a new structural motif, the winged helix structure. It is related to the well known helix turn helix motif, but defines a distinct nucleic acid recognition domain on its own. The structural similarity of the winged helix domain to other proteins (e.g., histone H5) is astonishing, since the degree of primary sequence similarity is not significant. This finding is a challenge to current structural predictions based only on primary sequences and it implies the involvement of this class of proteins in formerly unrelated processes. Thus, the observation of similar structures for histone H5 and the HNF-3γ fork head domain requires a functional analysis of fork head proteins in the process of nucleosome assembly and positioning (McPherson et al., 1993).

In vitro DNA binding site selection studies have shown that recognition is due to a conserved 7 bp motif, whose flanking domains differ by sequence and confer binding specificity. Presently it is not clear how this specification process is accomplished; it is possible that higher order structures (bending), charge patterns or more complicated structures may play a role. In this respect, it has also to be solved how differential affinities for specific targets are conferred by amino acid residues of the fork head domain which are not reported to be in contact with DNA. Furthermore, the allocation of a signal for nuclear targeting within this domain points towards other roles than nucleic acid recognition. Also, the function of the regions flanking the fork head domain has not been studied as extensively. The exploration of the transactivation process with respect to potential protein–protein contacts and post-translational modifications will be of major interest in future biochemical studies.

The role of winged helix proteins as regulatory keys in

embryogenesis and differentiated tissues has been a major field of interest during recent years. Many reports that have been reviewed here clearly demonstrate that this class of transcription factors serves important functions in the establishment of the body axis and in the differentiation and specification of tissues in all multicellular organisms. *Drosophila* mutants (fkh/croc), ectopic overexpression in amphibians (pintallavis) and knock-out experiments in mouse (HNF-3β/BF-1) have proven powerful tools to substantiate the function of individual fork head genes in embryogenesis. In many cases, it is already evident that the maintenance of a cellular differentiation state requires individual fork head factors. Future studies will render additional information for the regulatory capacity of winged helix proteins. It will also be explored in more detail, how these factors participate in embryonic induction and differentiation processes and how they are integrated into the regulatory cascades of transcription factors leading to the differentiation of specific tissues. In other words, we have to investigate regulatory mechanisms which govern expression of different members of this gene family and, simultaneously, identify target genes which are activated or repressed by individual winged helix proteins.

Involvement of fork head proteins in viral regulation and tumorigenesis illustrates that these genes do also act as proto-oncogenes. Probably, additional types of tumours will be characterized which are caused by dysregulation or mutations of specific fork head genes. Thus, after 5 years of extensive research on fork head proteins one has to conclude that this class of proteins is indeed an important 'characteristic DNA binding motif' as already foreseen by Weigel and Jäckle in 1990.

Acknowledgements

Our own investigations included in this review were supported by grants from the Deutsche Forschungsgemeinschaft (Kn 200/4-1) and Fonds der Chemischen Industrie.

Note added in proof

When this manuscript was already at proof stage we noticed an additional report [Dirksen and Jamrich (1995) Dev. Genet. 17, 107–116] on four *Xenopus* (XFKH3, 4, 5 and 6) and two zebrafish (zf-FKH1 and 2) genes. XFKH3 is identical with XFD-9, XFKH6 with XFD-6.

References

[1] Ang, S.L. and Rossant, J. (1994) HNF-3 beta is essential for node and notochord formation in mouse development. Cell 78, 561–574.

[2] Ang, S.L., Wierda, A., Wong, D., Stevens, K.A., Cascio, S., Rossant, J. and Zaret, K.S. (1993) The formation and maintenance of the definitive endoderm lineage in the mouse: involvement of HNF3/ forkhead proteins. Development 119, 1301–1315.

[3] Angrand, P.O., Coffinier, C. and Weiss, M.C. (1994) Response of the phosphoenolpyruvate carboxykinase gene to glucocorticoids depends on the integrity of the cAMP pathway. Cell. Growth Differ. 5, 957–966.

[4] Auge-Gouillou, C., Petropoulos, I. and Zakin, M.M. (1993) Liver-enriched HNF-3 alpha and ubiquitous factors interact with the human transferrin gene enhancer. FEBS Lett. 323, 4–10.

[5] Avraham, K.B., Fletcher, C., Overdier, D.G., Clevidence, D.E., Lai, E., Costa, R.H., Jenkins, N.A. and Copeland, N.G. (1995) Murine chromosomal location of eight members of the hepatocyte nuclear factor 3/fork head winged helix family of transcription factors. Genomics 25, 388–393.

[6] Avraham, K.B., Prezioso, V.R., Chen, W.S., Lai, E., Sladek, F.M., Zhong, W., Darnell, Jr., J.E., Jenkins, N.A. and Copeland, N.G. (1992) Murine chromosomal location of four hepatocyte-enriched transcription factors: HNF-3 alpha, HNF-3 beta, HNF-3 gamma, and HNF-4. Genomics 13, 264–248.

[7] Bassel-Duby, R., Hernandez, M.D., Yang, Q., Rochelle, J.M., Seldin, M.F. and Williams, R.S. (1994) Myocyte nuclear factor, a novel winged-helix transcription factor under both developmental and neural regulation in striated myocytes. Mol. Cell. Biol. 14, 4596–4605.

[8] Benit, P., Chanet, R., Fabre, F., Faye, G., Fukuhara, H. and Sor, F. (1992) Sequence of the sup61-RAD18 region on chromosome III of *Saccharomyces cerevisiae*. Yeast 8, 147–153.

[9] Bingle, C.D. and Gitlin, J.D. (1993) Identification of hepatocyte nuclear factor-3 binding sites in the Clara cell secretory protein gene. Biochem. J. 295, 227–232.

[10] Bingle, C.D., Hackett, B.P., Moxley, M., Longmore, W. and Gitlin, J.D. (1995) Role of hepatocyte nuclear factor-3 alpha and hepatocyte nuclear factor-3 beta in Clara cell secretory protein gene expression in the bronchiolar epithelium. Biochem. J. 308, 197–202.

[11] Bohinski, R.J., Di Lauro, R. and Whitsett, J.A. (1994) The lung-specific surfactant protein B gene promoter is a target for thyroid transcription factor 1 and hepatocyte nuclear factor 3, indicating common factors for organ-specific gene expression along the foregut axis. Mol. Cell. Biol. 14, 5671–5681.

[12] Bolce, M.E., Hemmati-Brivanlou, A. and Harland, R.M. (1993) XFKH2, a *Xenopus* HNF-3 alpha homologue, exhibits both activin-inducible and autonomous phases of expression in early embryos. Dev. Biol. 160, 413–423.

[13] Bork, P., Ouzounis, C., Sander, C., Scharf, M., Schneider, R. and Sonnhammer, E. (1992) Comprehensive sequence analysis of the 182 predicted open reading frames of yeast chromosome III. Protein Sci. 1, 1677–1690.

[14] Brennan, R.G. (1993) The winged-helix DNA-binding motif: another helix-turn-helix takeoff. Cell 74, 773–776.

[15] Cadigan, K.M., Grossniklaus, U. and Gehring, W.J. (1994a) Functional redundancy: the respective roles of the two sloppy paired genes in *Drosophila* segmentation. Proc. Natl. Acad. Sci. USA 91, 6324–6328.

[16] Cadigan, K.M., Grossniklaus, U. and Gehring, W.J. (1994b) Localized expression of sloppy paired protein maintains the polarity of *Drosophila* parasegments. Genes Dev. 8, 899–913.

[17] Cardinaux, J.R., Chapel, S. and Wahli, W. (1994) Complex organization of CTF/NF-I, C/EBP, and HNF3 binding sites within the promoter of the liver-specific vitellogenin gene. J. Biol. Chem. 269, 32947–32956.

[18] Chen, M., Hieng, S., Qian, X., Costa, R. and Ou, J.H. (1994) Regulation of hepatitis B virus ENI enhancer activity by hepatocyte-enriched transcription factor HNF3. Virology 205, 127–132.

[19] Cho, K.W., Blumberg, B., Steinbeisser, H. and De Robertis, E.M. (1991) Molecular nature of Spemann's organizer: the role

of the *Xenopus* homeobox gene goosecoid. Cell 67, 1111–1120.

[20] Clark, K.L., Halay, E.D., Lai, E. and Burley, S.K. (1993) Co-crystal structure of the HNF-3/fork head DNA-recognition motif resembles histone H5. Nature 364, 412–420.

[21] Clement, J.H., Fettes, P., Knöchel, S., Lef, J. and Knöchel, W. (1995) Bone morphogenetic protein 2 in the early development of *Xenopus laevis*. Mech. Dev. 52, 357–370.

[22] Clevidence, D.E., Overdier, D.G., Peterson, R.S., Porcella, A., Ye, H., Paulson, K.E. and Costa, R.H. (1994) Members of the HNF-3/forkhead family of transcription factors exhibit distinct cellular expression patterns in lung and regulate the surfactant protein B promoter. Dev. Biol. 1664, 195–209.

[23] Clevidence, D.E., Overdier, D.G., Tao, W., Qian, X., Pani, L., Lai, E. and Costa, R.H. (1993) Identification of nine tissue-specific transcription factors of the hepatocyte nuclear factor 3/forkhead DNA-binding-domain family. Proc. Natl. Acad. Sci. USA 90, 3948–3952.

[24] Cockell, M., Stolarczyk, D., Frutiger, S., Hughes, G.J., Hagenbüchle, O. and Wellauer, P.K. (1995) Binding sites for hepatocyte nuclear factor 3 beta or 3 gamma and pancreas transcription factor 1 are required for efficient expression of the gene encoding pancreatic alpha-amylase. Mol. Cell. Biol. 15, 1933–1941.

[25] Costa, R.H., Grayson, D.R. and Darnell, Jr., J.E. (1989) Multiple hepatocyte-enriched nuclear factors function in the regulation of transthyretin and alpha 1-antitrypsin genes. Mol. Cell. Biol. 9, 1415–1425.

[26] Dirksen, M.L. and Jamrich, M. (1992) A novel, activin-inducible, blastopore lip-specific gene of *Xenopus laevis* contains a fork head DNA-binding domain. Genes Dev. 6, 599–608.

[27] Donaldson, L.W., Petersen, J.M., Graves, B.J. and McIntosh, L.P. (1996) Solution structure of the ETS domain from murine Ets-1: a winged helix-turn-helix DNA binding motif. EMBO J. 15, 125–134.

[28] Downing, J.R., Khandekar, A., Shurtleff, S.A., Head, D.R., Parham, D.M., Webber, B.L., Pappo, A.S., Hulshof, M.G., Conn, W.P. and Shapiro, D.N. (1995) Multiplex RT-PCR assay for the differential diagnosis of alveolar rhabdomyosarcoma and Ewing's sarcoma. Am. J. Pathol. 146, 626–634.

[29] Drewes, T., Klein-Hitpass, L. and Ryffel, G.U. (1991) Liver specific transcription factors of the HNF3-, C/EBP- and LFB1-families interact with the A-activator binding site. Nucleic Acids Res. 19, 6383–6389.

[30] Echelard, Y., Epstein, D.J., St-Jacques, B., Shen, L., Mohler, J., McMahon, J.A. and McMahon, A.P. (1993) Sonic hedgehog, a member of a family of putative signaling molecules, is implicated in the regulation of CNS polarity. Cell 75, 1417–1430.

[31] Eguchi, H., Westin, S., Strom, A., Gustafsson, J.A. and Zaphiropoulos, P.G. (1991) Gene structure and expression of the rat cytochrome P450IIC13, a polymorphic, male-specific cytochrome in the P450IIC subfamily. Biochemistry 30, 10844–10849.

[32] Enerbäck, S., Ohlsson, B.G., Samuelsson, L. and Bjursell, G. (1992) Characterization of the human lipoprotein lipase (LPL) promoter: evidence of two cis-regulatory regions, LP-alpha and LP-beta, of importance for the differentiation-linked induction of the LPL gene during adipogenesis. Mol. Cell. Biol. 12, 4622–4633.

[33] Epstein, J.A., Lam, P., Jepeal, L., Maas, R.L. and Shapiro, D.N. (1995) Pax3 inhibits myogenic differentiation of cultured myoblast cells. J. Biol. Chem. 270, 11719–11722.

[34] Espinas, M.L., Roux, J., Ghysdael, J., Pictet, R. and Grange, T. (1994) Participation of Ets transcription factors in the glucocorticoid response of the rat tyrosine aminotransferase gene. Mol. Cell. Biol. 14, 4116–4125.

[35] Espinas, M.L., Roux, J., Pictet, R. and Grange, T. (1995) Glucocorticoids and protein kinase A coordinately modulate transcription factor recruitment at a glucocorticoid-responsive unit. Mol. Cell. Biol. 15, 5346–5354.

[36] Fahrner, J., Labruyere, W.T., Gaunitz, C., Moorman, A.F., Gebhardt, R. and Lamers, W.H. (1993) Identification and functional characterization of regulatory elements of the glutamine synthetase gene from rat liver. Eur. J. Biochem. 213, 1067–1073.

[37] Fainsod, A., Steinbeisser, H. and De Robertis, E.M. (1994) On the function of BMP-4 in patterning the marginal zone of the *Xenopus* embryo. EMBO J. 1354, 5015–5025.

[38] Fogh, R.H., Ottleben, G., Ruterjans, H., Schnarr, M., Boelens, R. and Kaptein, R. (1994) Solution structure of the LexA repressor DNA binding domain determined by 1H NMR spectroscopy. EMBO J. 13, 3936–3944.

[39] Fredericks, W.J., Galili, N., Mukhopadhyay, S., Rovera, G., Bennicelli, J., Barr, F.G. and Rauscher, Jr., F. (1995) The PAX3-FKHR fusion protein created by the t(2;13) translocation in alveolar rhabdomyosarcomas is a more potent transcriptional activator than PAX3. Mol. Cell. Biol. 15, 1522–1535.

[40] Galili, N., Davis, R.J., Fredericks, W.J., Mukhopadhyay, S., Rauscher, F.J.D., Emanuel, B.S., Rovera, G. and Barr, F.G. (1993) Fusion of a fork head domain gene to PAX3 in the solid tumour alveolar rhabdomyosarcoma. Nature Genet. 5, 230–235.

[41] Gaul, U. and Weigel, D. (1990) Regulation of Krüppel expression in the anlage of the Malpighian tubules in the *Drosophila* embryo. Mech. Dev. 33, 57–67.

[42] Gregori, C., Kahn, A. and Pichard, A.L. (1993) Competition between transcription factors HNF1 and HNF3, and alternative cell-specific activation by DBP and C/EBP contribute to the regulation of the liver-specific aldolase B promoter. Nucleic Acids Res. 21, 897–903.

[43] Gregori, C., Kahn, A. and Pichard, A.L. (1994) Activity of the rat liver-specific aldolase B promoter is restrained by HNF3. Nucleic Acids Res. 22, 1242–1246.

[44] Grossniklaus, U., Pearson, R.K. and Gehring, W.J. (1992) The *Drosophila* sloppy paired locus encodes two proteins involved in segmentation that show homology to mammalian transcription factors. Genes Dev. 6, 1030–1051.

[45] Häcker, U., Grossniklaus, U., Gehring, W.J. and Jäckle, H. (1992) Developmentally regulated *Drosophila* gene family encoding the fork head domain. Proc. Natl. Acad. Sci. USA 89, 8754–8758.

[46] Häcker, U., Kaufmann, E., Hartmann, C., Jürgens, G., Knöchel, W. and Jäckle, H. (1995) The *Drosophila* fork head domain protein crocodile is required for the establishment of head structures. EMBO J. 14, 5306–5317.

[47] Hackett, B.P., Brody, S.L., Liang, M., Zeitz, I.D., Bruns, L.A. and Gitlin, J.D. (1995) Primary structure of hepatocyte nuclear factor/forkhead homologue 4 and characterization of gene expression in the developing respiratory and reproductive epithelium. Proc. Natl. Acad. Sci. USA 92, 4249–4253.

[48] Hatini, V., Tao, W. and Lai, E. (1994) Expression of winged helix genes, BF-1 and BF-2, define adjacent domains within the developing forebrain and retina. J. Neurobiol. 25, 1293–1309.

[49] Hermann-Le Denmat, S., Werner, M., Sentenac, A. and Thuriaux, P. (1994) Suppression of yeast RNA polymerase III mutations by FHL1, a gene coding for a fork head protein involved in rRNA processing. Mol. Cell. Biol. 14, 2905–2913.

[50] Hoch, M., Broadie, K., Jäckle, H. and Skaer, H. (1994) Sequential fates in a single cell are established by the neurogenic cascade in the Malpighian tubules of *Drosophila*. Development 120, 33439–3450.

[51] Hope, I.A. (1994) PES-1 is expressed during early embryogenesis in *Caenorhabditis elegans* and has homology to the fork head family of transcription factors. Development 120, 505–514.

[52] Hromas, R., Klemsz, M., Amaravadi, L., Hufford, T., Huang, I., Desai, A., Srour, E., Bruno, E. and Hoffman, R. (1994) *Drosophila* forkhead homologues are expressed in CD34+/HLA-DR-

primitive human hematopoietic progenitors. Leuk. Lymphoma 15, 439–444.

[53] Hromas, R., Moore, J., Johnston, T., Socha, C. and Klemsz, M. (1993) *Drosophila* forkhead homologues are expressed in a lineage-restricted manner in human hematopoietic cells. Blood 81, 2854–2859.

[54] Ip, Y.T., Poon, D., Stone, D., Granner, D.K. and Chalkley, R. (1990) Interaction of a liver-specific factor with an enhancer 4.8 kilobases upstream of the phosphoenolpyruvate carboxykinase gene. Mol. Cell. Biol. 10, 3770–3781.

[55] Jackson, D.A., Rowader, K.E., Stevens, K., Jiang, C., Milos, P. and Zaret, K.S. (1993) Modulation of liver-specific transcription by interactions between hepatocyte nuclear factor 3 and nuclear factor 1 binding DNA in close apposition. Mol. Cell. Biol. 13, 2401–2410.

[56] Kaestner, K.H., Lee, K.H., Schlöndorff, J., Hiemisch, H., Monaghan, A.P. and Schütz, G. (1993) Six members of the mouse forkhead gene family are developmentally regulated. Proc. Natl. Acad. Sci. USA 90, 7628–7631.

[57] Kaestner, K.H., Hiemisch, H., Luckow, B. and Schütz, G. (1994) The HNF-3 gene family of transcription factors in mice: gene structure, cDNA sequence, and mRNA distribution. Genomics 20, 377–385.

[58] Kaestner, K.H., Monaghan, A.P., Kern, H., Ang, S.-L., Weitz, S., Lichter, P. and Schütz, G. (1995) The mouse fkh-2 gene. J. Biol. Chem. 270, 30029–30035

[59] Kaufmann, E., Hoch, M. and Jäckle, H. (1994) The interaction of DNA with the DNA-binding domain encoded by the *Drosophila* gene fork head. Eur. J. Biochem. 223, 329–337.

[60] Kaufmann, E., Müller, D. and Knöchel, W. (1995) DNA recognition site analysis of *Xenopus* winged helix proteins. J. Mol. Biol. 248, 239–254.

[61] King, M.W. and Moore, M.J. (1994) Novel Hox, POU and FKH genes expressed during bFGF-induced mesodermal differentiation in *Xenopus*. Nucleic Acids Res. 22, 3990–3996.

[62] Knöchel, S., Lef, J., Clement, J., Klocke, B., Hille, S., Köster, M. and Knöchel, W. (1992) Activin A induced expression of a fork head related gene in posterior chordamesoderm (notochord) of *Xenopus laevis* embryos. Mech. Dev. 38, 157–165.

[63] Kuo, C.J., Conley, P.B., Chen, L., Sladek, F.M., Darnell, Jr., J.E., and Crabtree, G.R. (1992) A transcriptional hierarchy involved in mammalian cell-type specification. Nature 355, 457–461.

[64] Kuzin, B., Tillib, S., Sedkov, Y., Mizrokhi, L. and Mazo, A. (1994) The *Drosophila* trithorax gene encodes a chromosomal protein and directly regulates the region-specific homeotic gene fork head. Genes Dev. 8424, 2478–2490.

[65] Lai, E., Clark, K.L., Burley, S.K. and Darnell, Jr., J.E., (1993) Hepatocyte nuclear factor 3/fork head or 'winged helix' proteins: a family of transcription factors of diverse biologic function. Proc. Natl. Acad. Sci. USA 90, 10421–10423.

[66] Lai, E., Prezioso, V.R., Smith, E., Litvin, O., Costa, R.H. and Darnell, Jr., J.E., (1990) HNF-3A, a hepatocyte-enriched transcription factor of novel structure is regulated transcriptionally. Genes Dev. 4, 1427–1436.

[67] Lai, E., Prezioso, V.R., Tao, W.F., Chen, W.S. and Darnell, Jr., J.E., (1991) Hepatocyte nuclear factor 3 alpha belongs to a gene family in mammals that is homologous to the *Drosophila* homeotic gene fork head. Genes Dev. 5, 416–427.

[68] Lef, J., Clement, J.H., Oschwald, R., Köster, M. and Knöchel, W. (1994) Spatial and temporal transcription patterns of the forkhead related XFD- 2/XFD-2' genes in *Xenopus laevis* embryos. Mech. Dev. 45, 117–26.

[69] Lef, J., Dege, P., Scheucher, M., Forsbach-Birk, V., Clement, J.H. and Knöchel, W. (1996) A fork head related multigene family is transcribed in *Xenopus laevis* embryos. Int. J. Dev. Biol., in press.

[70] Li, C. and Tucker, P.W. (1993) DNA-binding properties and secondary structural model of the hepatocyte nuclear factor 3/fork head domain. Proc. Natl. Acad. Sci. USA 90, 11583–11587.

[71] Li, J. and Vogt, P.K. (1993) The retroviral oncogene qin belongs to the transcription factor family that includes the homeotic gene fork head. Proc. Natl. Acad. Sci. USA 90, 4490–4494.

[72] Li, C., Lai, C.F., Sigman, D.S. and Gaynor, R.B. (1991) Cloning of a cellular factor, interleukin binding factor, that binds to NFAT-like motifs in the human immunodeficiency virus long terminal repeat. Proc. Natl. Acad. Sci. USA 88, 7739–7743.

[73] Li, C., Lusis, A.J., Sparkes, R., Nirula, A. and Gaynor, R. (1992a) Characterization and chromosomal mapping of the gene encoding the cellular DNA binding protein ILF. Genomics 13, 665–671.

[74] Li, C., Lusis, A.J., Sparkes, R., Tran, S.M. and Gaynor, R. (1992b) Characterization and chromosomal mapping of the gene encoding the cellular DNA binding protein HTLF. Genomics 13, 658–664.

[75] Liang, H., Olejniczak, E.T., Mao, X., Nettesheim, D.G., Yu, L., Thompson, C.B. and Fesik, S.W. (1994) The secondary structure of the ets domain of human Fli-1 resembles that of the helix-turn-helix DNA-binding motif of the *Escherichia coli* catabolite gene activator protein. Proc. Natl. Acad. Sci. USA 91, 11655–11659.

[76] Liu, C., Mason, W.S. and Burch, J.B. (1994) Identification of factor-binding sites in the duck hepatitis B virus enhancer and in vivo effects of enhancer mutations. J. Virol. 68, 2286–2296.

[77] Macdonald, R., Xu, Q., Barth, K.A., Mikkola, I., Holder, N., Fjose, A., Krauss, S. and Wilson, S.W. (1994) Regulatory gene expression boundaries demarcate sites of neuronal differentiation in the embryonic zebrafish forebrain. Neuron 1327, 1039–1053.

[78] Mach, V., Takiya, S., Ohno, K., Handa, H., Imai, T. and Suzuki, Y. (1995) Silk gland factor-1 involved in the regulation of *Bombyx* sericin-1 gene contains fork head motif. J. Biol. Chem. 270, 9340–9346.

[79] Macleod, K., Leprince, D. and Stehelin, D. (1992) The ets gene family. Trends Biochem. Sci. 17, 251–256.

[80] McMahon, A.P. (1994) Mouse development: winged-helix in axial patterning. Curr. Biol. 4, 903–906.

[81] McPherson, C.E., Shim, E.Y., Friedman, D.S. and Zaret, K.S. (1993) An active tissue-specific enhancer and bound transcription factors existing in a precisely positioned nucleosomal array. Cell 75, 387–398.

[82] Miller, L.M., Gallegos, M.E., Morisseau, B.A. and Kim, S.K. (1993) lin-31, a *Caenorhabditis elegans* HNF-3/fork head transcription factor homolog, specifies three alternative cell fates in vulval development. Genes Dev. 7, 933–947.

[83] Millonig, J.H., Emerson, J.A., Levorse, J.M. and Tilghman, S.M. (1995) Molecular analysis of the distal enhancer of the mouse alpha-fetoprotein gene. Mol. Cell. Biol. 15, 3848–3856.

[84] Mitchell, P.J. and Tjian, R. (1989) Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. Science 245, 371–378.

[85] Miura, N., Wanaka, A., Tohyama, M. and Tanaka, K. (1993) MFH-1, a new member of the fork head domain family, is expressed in developing mesenchyme. FEBS Lett. 326, 171–176.

[86] Monaghan, A.P., Kaestner, K.H., Grau, E. and Schütz, G. (1993) Postimplantation expression patterns indicate a role for the mouse forkhead/HNF-3 alpha, beta and gamma genes in determination of the definitive endoderm, chordamesoderm and neuroectoderm. Development 119, 567–578.

[87] Murphy, D.B., Wiese, S., Burfeind, P., Schmundt, D., Mattei, M.G., Schulz-Schaeffer, W. and Thies, U. (1994) Human brain factor 1, a new member of the fork head gene family. Genomics 2124, 551–557.

[88] Nehls, M., Pfeifer, D., Schorpp, M., Hedrich, H. and Boenm, T. (1994) New member of the winged-helix protein family disrupted in mouse and rat nude mutations. Nature 372, 103–107.

[89] Nelson, H.C.M. (1995) Structure and function of DNA-binding proteins. Curr. Opin. Genet. Dev. 5, 180–189.

[90] Nitsch, D. and Schütz, G. (1993) The distal enhancer implicated in the developmental regulation of the tyrosine aminotransferase gene is bound by liver-specific and ubiquitous factors. Mol. Cell. Biol. 13, 4494–504.

[91] Nitsch, D., Boshart, M. and Schütz, G. (1993a) Activation of the tyrosine aminotransferase gene is dependent on synergy between liver-specific and hormone-responsive elements. Proc. Natl. Acad. Sci. USA 90, 5479–5483.

[92] Nitsch, D., Boshart, M. and Schütz, G. (1993b) Extinction of tyrosine aminotransferase gene activity in somatic cell hybrids involves modification and loss of several essential transcriptional activators. Genes Dev. 7, 308–319.

[93] Ori, A. and Shaul, Y. (1995) Hepatitis B virus enhancer binds and is activated by the hepatocyte nuclear factor 3. Virology 207, 98–106.

[94] Overdier, D.G., Porcella, A. and Costa, R.H. (1994) The DNA-binding specificity of the hepatocyte nuclear factor 3/forkhead domain is influenced by amino-acid residues adjacent to the recognition helix. Mol. Cell. Biol. 14, 2755–2766.

[95] Pani, L., Overdier, D.G., Porcella, A., Qian, X., Lai, E. and Costa, R.H. (1992a) Hepatocyte nuclear factor 3β contains two transcriptional activation domains, one of which is novel and conserved with the *Drosophila* fork head protein. Mol. Cell. Biol. 12, 3723–3732.

[96] Pani, L., Quian, X.B., Clevidence, D. and Costa, R.H. (1992b) The restricted promoter activity of the liver transcription factor hepatocyte nuclear factor 3 beta involves a cell-specific factor and positive autoactivation. Mol. Cell. Biol. 12, 552–562.

[97] Parry, P., Wei, Y. and Evans, G. (1994) Cloning and characterization of the t(X;11) breakpoint from a leukemic cell line identify a new member of the forkhead gene family. Genes Chromosom. Cancer 11, 79–84.

[98] Pierrou, S., Hellqvist, M., Samuelsson, L., Enerbäck, S. and Carlsson, P. (1994) Cloning and characterization of seven human forkhead proteins: binding site specificity and DNA bending. EMBO J. 1354, 5002–5012.

[99] Ptashne, M. (1988) How eukaryotic transcriptional activators work. Nature 335, 683–689.

[100] Ptashne, M. and Gann, A.A. (1990) Activators and targets. Nature 346, 329–331.

[101] Qian, X. and Costa, R.H. (1995) Analysis of hepatocyte nuclear factor-3 beta protein domains required for transcriptional activation and nuclear targeting. Nucleic Acids Res. 23, 1184–1191.

[102] Qian, X., Samadani, U., Porcella, A. and Costa, R.H. (1995) Decreased expression of hepatocyte nuclear factor 3 alpha during the acute-phase response influences transthyretin gene transcription. Mol. Cell. Biol. 15, 1364–1376.

[103] Ramakrishnan, V., Finch, J.T., Graziano, V., Lee, P.L. and Sweet, R.M. (1993) Crystal structure of globular domain of histone H5 and its implications for nucleosome binding. Nature 362, 219–223.

[104] Raney, A.K., Zhang, P. and McLachlan, A. (1995) Regulation of transcription from the hepatitis B virus large surface antigen promoter by hepatocyte nuclear factor 3. J. Virol. 69, 3265–3272.

[105] Raymondjean, M., Pichard, A.L., Gregori, C., Ginot, F. and Kahn, A. (1991) Interplay of an original combination of factors: C/EBP, NFY, HNF3, and HNF1 in the rat aldolase B gene promoter. Nucleic Acids Res. 19, 6145–6153.

[106] Riddle, R.D., Johnson, R.L., Laufer, E. and Tabin, C. (1993) Sonic hedgehog mediates the polarizing activity of the ZPA. Cell 75, 1401–1416.

[107] Roux, J., Pictet, R. and Grange, T. (1995) Hepatocyte nuclear factor 3 determines the amplitude of the glucocorticoid response of the rat tyrosine aminotransferase gene. DNA Cell. Biol. 14, 385–396.

[108] Ruiz i Altaba, A. and Jessell, T.M. (1992) Pintallavis, a gene expressed in the organizer and midline cells of frog embryos: involvement in the development of the neural axis. Development 116, 81–93.

[109] Ruiz i Altaba, A.R., Cox, C., Jessell, T.M. and Klar, A. (1993a) Ectopic neural expression of a floor plate marker in frog embryos injected with the midline transcription factor Pintallavis. Proc. Natl. Acad. Sci. USA 90, 8268–72.

[110] Ruiz i Altaba, A., Prezioso, V.R., Darnell, J.E. and Jessell, T.M. (1993b) Sequential expression of HNF-3 beta and HNF-3 alpha by embryonic organizing centers: the dorsal lip/node, notochord and floor plate. Mech. Dev. 44, 91–108.

[111] Sasaki, H. and Hogan, B.L. (1993) Differential expression of multiple fork head related genes during gastrulation and axial pattern formation in the mouse embryo. Development 118, 47–59.

[112] Sasaki, H. and Hogan, B.L. (1994) HNF-3 beta as a regulator of floor plate development. Cell 76, 103–115.

[113] Sawaya, P.L. and Luse, D.S. (1994) Two members of the HNF-3 family have opposite effects on a lung transcriptional element: HNF-3 alpha stimulates and HNF-3 beta inhibits activity of region I from the Clara cell secretory protein (CCSP) promoter. J. Biol. Chem. 269, 22211–22216.

[114] Sawaya, P.L., Stripp, B.R., Whitsett, J.A. and Luse, D.S. (1993) The lung-specific CC10 gene is regulated by transcription factors from the AP-1, octamer, and hepatocyte nuclear factor 3 families. Mol. Cell. Biol. 13, 3860–3871.

[115] Scheucher, M., Dege, P., Lef, J., Hille, S. and Knöchel, W. (1995) Transcription patterns of four different fork head/HNF-3 related genes (XFD-4, 6, 9 and 10) in *Xenopus laevis* embryos. Rouxs Arch. Dev. Biol. 204, 203–211.

[116] Schultz, S.C., Shields, G.C. and Steitz, T.A. (1991) Crystal structure of a CAP-DNA complex: the DNA is bent by 90 degrees. Science 253, 1001–1007.

[117] Shapiro, D.N., Sublett, J.E., Li, B., Downing, J.R. and Naeve, C.W. (1993) Fusion of PAX3 to a member of the forkhead family of transcription factors in human alveolar rhabdomyosarcoma. Cancer Res. 53, 5108–5112.

[118] Shaw, P.M., Weiss, M.C. and Adesnik, M. (1994) Hepatocyte nuclear factor 3 is a major determinant of CYP2C6 promoter activity in hepatoma cells. Mol. Pharmacol. 46, 79–87.

[119] Strähle, U., Blader, P., Henrique, D. and Ingham, P.W. (1993) Axial, a zebrafish gene expressed along the developing body axis, shows altered expression in cyclops mutant embryos. Genes Dev. 7, 1436–1446.

[120] Sublett, J.E., Jeon, I.S. and Shapiro, D.N. (1995) The alveolar rhabdomyosarcoma PAX3/FKHR fusion protein is a transcriptional activator. Oncogene 11, 545–552.

[121] Tao, W. and Lai, E. (1992) Telencephalon-restricted expression of BF-1, a new member of the HNF-3/fork head gene family, in the developing rat brain. Neuron 8, 957–966.

[122] Tepass, U., Pessler, L.I., Aziz, A. and Hartenstein, V. (1994) Embryonic origin of hemocytes and their relationship to cell death in *Drosophila*. Development 120, 1829–1837.

[123] Tjian, R. and Maniatis, T. (1994) Transcriptional activation: a complex puzzle with few easy pieces. Cell 77, 5–8.

[124] Tremp, G.L., Duchange, N., Branellec, D., Cereghini, S., Tailleux, A., Berthou, L., Fievet, C., Touchet, N., Schombert, B., Fruchart, A., Zakin, M.M. and Deneflc, P. (1995) A 700-bp fragment of the human antithrombin III promoter is sufficient to confer high, tissue-specific expression on human apolipoprotein A-II in transgenic mice. Gene 1565, 199–205.

[125] Ueguki, K., Shirakawa, M., Harada, H., Taniguchi, T. and Kyo-

goku, Y. (1995) Secondary structure and folding topology of the DNA binding domain of interferon regulatory factor 2, as revealed by NMR spectroscopy. FEBS Lett. 359, 184–188.

[126] Vuister, G.W., Kim, S.J., Wu, C. and Bax, A. (1994) NMR evidence for similarities between the DNA-binding regions of *Drosophila melanogaster* heat shock factor and the helix-turn-helix and HNF-3/forkhead families of transcription factors. Biochemistry 33, 10–16.

[127] Weigel, D. and Jäckle, H. (1990) The fork head domain: a novel DNA binding motif of eukaryotic transcription factors. Cell 63, 455–456.

[128] Weigel, D., Jürgens, G., Küttner, F., Seifert, E. and Jäckle, H. (1989) The homeotic gene fork head encodes a nuclear protein and is expressed in the terminal regions of the *Drosophila* embryo. Cell 57, 645–658.

[129] Weigel, D. Jurgens, G., Klingler, M. and Jäckle, H. (1990a) Two gap genes mediate maternal terminal pattern information in *Drosophila*. Science 248, 495–498.

[130] Weigel, D., Seifert, E., Reuter, D. and Jäckle, H. (1990b) Regulatory elements controlling expression of the *Drosophila* homeotic gene fork head. EMBO J. 9, 1199–1207.

[131] Weinstein, D.C., Ruiz i Altaba, A., Chen, W.S., Hoodless, P., Prezioso, V.R., Jessell, T.M. and Darnell, Jr., J.E., (1994) The winged-helix transcription factor HNF-3 beta is required for notochord development in the mouse embryo. Cell 78, 575–588.

[132] Wiese, S., Murphy, D.B., Schlung, A., Burfeind, P., Schmundt, D., Schnülle, V., Mattei, M.G. and Thies, U. (1995) The genes for human brain factor 1 and 2, members of the fork head gene family, are clustered on chromosome 14q. Biochim. Biophys. Acta 1262, 105–112.

[133] Wolberger, C., Vershon, A.K., Liu, B., Johnson, A.D. and Pabo, C.O. (1991) Crystal structure of a MAT alpha 2 homeodomain-operator complex suggests a general model for homeodomain-DNA interactions. Cell 67, 517–528.

[134] Xuan, S., Baptista, C.A., Balas, G., Tao, W., Soares, V.C. and Lai, E. (1995) Winged helix transcription factor BF-1 is essential for the development of the cerebral hemispheres. Neuron 14, 1141–1152.

[135] Zhao, J.J., Lazzarini, R.A. and Pick, L. (1993) The mouse Hox-1.3 gene is functionally equivalent to the *Drosophila* Sex combs reduced gene. Genes Dev. 7, 343–354.

[136] Zhu, G. and Davis, T.N. (1995) Two fork head homologs in *S. cerevisae*. EMBL data library (ID: SCFKH1A and SCFKH2A), submitted.

[137] Zhu, G., Müller, E.G., Amacher, S.L., Northrop, J.L. and Davis, T.N. (1993) A dosage-dependent suppressor of a temperature-sensitive calmodulin mutant encodes a protein related to the fork head family of DNA-binding proteins. Mol. Cell. Biol. 13, 1779–1787.